L1 ORF1p Is an Early Diagnostic Marker of Cancer and Its Precursor Syndromes

PhD Dissertation

Student: Khaldoon Sadiq Ahmed Abdullah, *BSc, Msc.*

Supervisor: Dr. Mates Lajos, *PhD*

Biological Research Centre

Doctoral School of Multidisciplinary Medical Sciences

Albert Szent-Györgyi Medical School

University of Szeged

Szeged, Hungary

2025

LIST OF PUBLICATIONS RELATED TO THE SUBJECT OF THE THESIS

I. Karkas R, *__Abdullah KSA__, Kaizer L, Ürmös Á, Raya M, Tiszlavicz L, Pankotai T, Nagy I, Mátés L, Sükösd F. L1 ORF1p is a Promising Biomarker in Cervical Intraepithelial Neoplasia Degree Assessment. Int J Gynecol Pathol. 2024 Jun 26. doi: 10.1097/PGP.0000000000001035. Epub ahead of print. PMID: 38920137. (Impact factor 3.32 (Q1)).

II. Imre G, Takács B, Czipa E, Drubi AB, Jaksa G, Latinovics D, Nagy A, Karkas R, Hudoba L, Vásárhelyi BM, Pankotai-Bodó G, Blastyák A, Hegedűs Z, Germán P, Bálint B, **Ahmed Abdullah KS**, Kopasz AG, Kovács A, Nagy LG, Sükösd F, Pintér L, Rülicke T, Barta E, Nagy I, Haracska L, Mátés L. Prolonged activity of the transposase helper may raise safety concerns during DNA transposon-based gene therapy. Mol Ther Methods Clin Dev. 2023 Mar 14;29:145-159. doi: 10.1016/j.omtm.2023.03.003. PMID: 37025950; PMCID: PMC10070360. *(Impact factor 5.84 (Q1))*

***Shared First Authorship**

LIST OF OTHER PUBLICATIONS DURING PhD STUDY

Full Papers:

I. Kopasz AG, Pusztai DZ, Karkas R, Hudoba L, **Abdullah KSA**, Imre G, Pankotai-Bodó G, Migh E, Nagy A, Kriston A, Germán P, Drubi AB, Molnár A, Fekete I, Dani VÉ, Ocsovszki I, Puskás LG, Horváth P, Sükösd F, Mátés L. A versatile transposon-based technology to generate loss- and gain-of-function phenotypes in the mouse liver. BMC Biol. 2022 Apr 1;20(1):74. doi: 10.1186/s12915-022-01262-x. PMID: 35361222; PMCID: PMC8974095 *(Impact factor 7.43 (Q1))*

Impact factor publications related to PhD thesis: 16.59

Presentations, Rewards and PATENTs:

I. **Abdullah KSA**, Réka Karkas, Pankotai-Bodó G, Sükösd F, Lajos Mates. L1 retrotransposition, a major character in human carcinogenesis? a study based on human tumor bank data. STRAUB Napok; 2022 May 25-27; Szeged, Hungary.

II. **Abdullah KSA**, Réka Karkas, Pankotai-Bodó G, Sükösd F, Lajos Mates. L1 retrotransposition, A Major Character in Human Carcinogensis. Membrane Transport Conference; 2022 May 17-19; Sümeg, Hungary.

III. Anna Georgina Kopasz, **Khaldoon S.Abdullah**, Dávid Pusztai, Réka Karkas, Liza Hudoba, Gergely Imre, Lajos Mates. A novel approach to silence p53 tumor suppressor gene in a somatically transgenic

mouse model. Proceedings of the 2nd V4SDB International Conference; 2021 September 2-5; Szeged, Hungary.

IV. Réka Karkas, András Blastyák, Liza Hudoba, Andrea Bakne Drubi, Gergely Imre, **Khaldoon S.Abdullah**, Lajos Mates. Comparison of the DNA transposon-based approaches in the gene therapy of the Mouse Model of Tyrosinemia Type l. Proceedings of the 2nd HMAA-hc International Conference; 2021 August 27-28; Balatonfured, Hungary.

V. Straub Young Scientist Prize 2022 , Szeged, Hungary 2022.

VI. Kopasz AG, Pusztai DZ, Karkas R, Hudoba L, Abdullah KSA, Imre G, Nagy A, Sükösd F, Mátés L. A versatile transposon-based technology to generate loss- and gain-of-function phenotypes in the mouse liver.2023; patent submission ID: P2200101.

VII. Karkas R, Abdullah KSA , Imre G, Nagy A, Sükösd F, Mátés. L Retrotransposon-based technique to examine the toxicity of chemicals and drugs in mouse liver. 2023; patent submission ID: P2200102.

LIST OF ABBREVIATIONS

APOBEC3: Apolipoprotein B mRNA Editing Catalytic Polypeptide 3

AUC: Area Under the Curve

bp: Base Pair

BRCA1: Breast Cancer 1

CI: Confidence Interval

CIN: Cervical Intraepithelial Neoplasia

CTD: C-terminal Domain

DCLRE1C: DNA Cross-Link Repair 1C

DNA: Deoxyribonucleic Acid

DNA-PKCs: DNA-Dependent Protein Kinase, Catalytic Subunit

ERV: Endogenous Retroviruses

FFPE: Formalin-Fixed Paraffin-Embedded

GAPDH: Glyceraldehyde 3-phosphate Dehydrogenase

hnRNPA1: Heterogeneous Nuclear Ribonucleoprotein A1

HPV: Human Papillomavirus

HSIL: High-grade Squamous Intraepithelial Lesion

IHC: Immunohistochemistry

KAP1/2: Kinase Associated Protein 1/2

Kb: Kilobases

kDa: Kilodalton

L1/L1: Long Interspersed Nuclear Element-1

L1ORF: L1 Open Reading Frames

LIG4: Ligase 4

LSIL: Low-grade Squamous Intraepithelial Lesion

LTR: Long Terminal Repeat

MEPC2: Methyl-CpG-binding Protein 2

MOV10: Moloney Leukemia Virus 10

mRNA: Messenger RNA

NHEJ: Non-Homologous End-Joining

NILM: Negative for Intraepithelial Lesion or Malignancy

NuRD: Nucleosomal and Remodeling Deacetylase

ORFs: Open Reading Frames

PCR: Polymerase Chain Reaction

qPCR: Quantitative Polymerase Chain Reaction

RISC: RNA-Induced Silencing Complex

RNA: Ribonucleic Acid

RNP: Ribonucleoprotein

RPA: Replication Protein A

RRM: RNA Recognition Motif

SCC: Squamous Cell Carcinoma

SINEs: Short Interspersed Nuclear Elements

SIRT6/7: Sirtuin 6/7

TEs: Transposable Elements

TNPO1: Transportin-1

TPRT: Target-Primed Reverse Transcription

TSDs: Target-Site Duplications

UTR: Untranslated Region

XRCC4/6: X-Ray Repair Cross Complementing 4/6

ZAP: Zinc Finger Antiviral Protein

Table of Contents

1. INTRODUCTION

1.1 Transposable elements

Transposable elements (TEs) also known as "hopping genes," are DNA sequences that can move in a genome, and thus play a significant role in genomic evolution and function. TEs were first figured out by Barbara McClintock in corn and have since been recognised across all forms of life, from bacteria to humans [1]. In the human genome, TEs comprise approximately 45-66% of the total DNA content [2,3]. These elements are categorized into two principal categories (Fig. 1): the DNA transposons, which constitute approximately 3% of the human genome, move with cut and paste mechanism [1]. In contrast, retrotransposons spread through an RNA intermediate in a "copy-and-paste" mechanism and are further classified into long terminal repeat (LTR) and non-LTR retrotransposons [1]. The most prevalent TEs in the human genome are the non-LTR retrotransposons, particularly the long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs), which collectively constitute approximately 21% and 13% of the genome, respectively [2]. However, many TEs are deactivated because of the accumulation of mutations over time, some remain active and can affect gene expression, cause genomic instability and contribute to genetic diversity [3]. Recent studies have identified TEs as playing a role in several biological processes, including embryonic development, immune response, and neurological functions [4, 5].

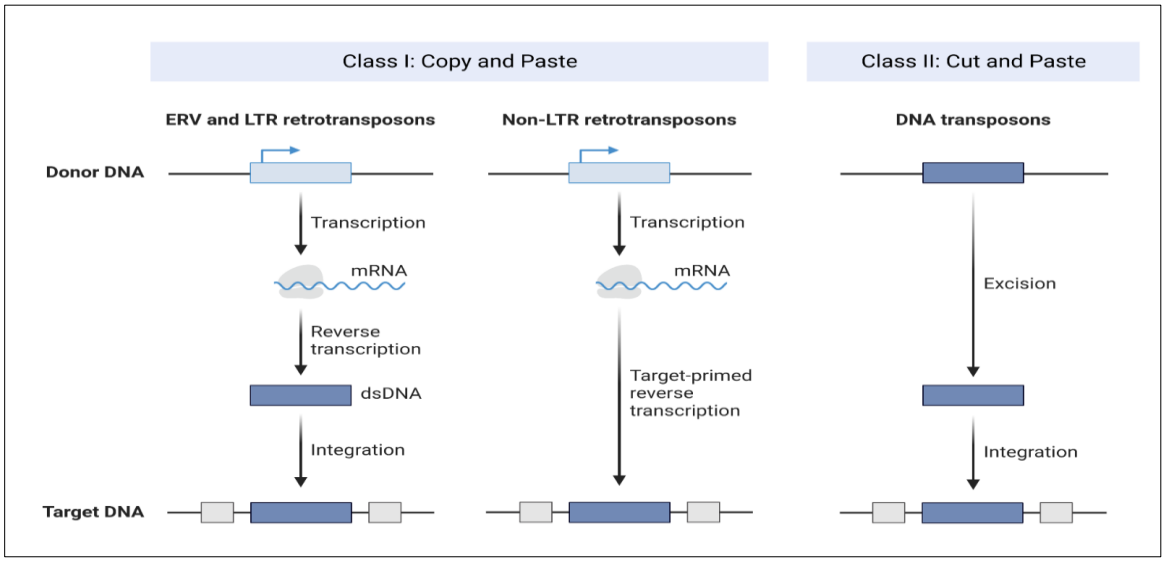


Figure 1. The main classes of mobile genetic elements. ERV, Endogenous Retroviruses, LTR, Long Terminal Repeats (Reproduced with Biorender.com).

1.2 Long Interspersed Nuclear Element-1 (LINE-1 or L1) retrotransposons

L1 retrotransposons are autonomous mobile genetic elements that cover approximately 17% of the human genome [6], stand for the most abundant class of transposable elements. Although the abundance of these sequences, only a modest percentage maintains the capacity to move, as the majority of L1 sequences are inactivated as a consequence of 5´ truncations, mutations, or rearrangements. Nevertheless, the activity of these elements has been linked to genomic instability, particularly in the context of cancer, where they have been observed to disrupt gene function and promote oncogenesis [6]. These elements play a pivotal role in genome plasticity and evolution, contributing to genetic diversity and potentially adaptive changes. Recent studies have demonstrated that L1s are involved in a number of physiological processes, including early embryonic development and neurogenesis. In embryonic stem cells and early embryos, L1 expression is typically elevated and appears to be crucial for normal development. This elevated activity may contribute to the creation of genetic mosaicism in the developing organism, thereby enhancing cellular diversity and function. In the brain, somatic L1 retrotransposition has been observed in neuronal progenitor cells, indicating a potential involvement in neuronal diversity and brain plasticity. It is possible that L1 mediated insertions in neuronal genomes may influence neuronal function and contribute to individual differences in brain function and behaviour. In addition to their involvement in development and neurogenesis, L1 elements have been linked to genome organisation and gene regulation. The insertion of L1 elements can affect gene expression through several mechanisms, including the introduction of new regulatory sequences, alteration of chromatin structure and the creation of new splice sites [7]. Recent research has demonstrated that L1 elements can function as alternative promoters for cellular genes, thereby expanding the diversity of transcripts and proteins produced from a single gene locus. Furthermore, L1-derived sequences have been demonstrated to facilitate the formation of topologically associating domains (TADs) in the genome, influencing three-dimensional chromatin organisation and gene regulation [8]. Notably, L1 elements have also been linked to DNA damage repair processes. Some studies have proposed that L1 endonuclease activity may play a role in resolving stalled replication forks and other forms of DNA damage, potentially contributing to genome stability under specific conditions [9].

1.3 Structure and function of L1 encoded proteins

The full-length mammalian L1 retrotransposons are approximately 6 kilobases in length. Two open reading frames (ORFs), which encode for L1ORF1p and L1ORF2p, are flanked by a ~900 bp 5' untranslated region (UTR) and a ~200 bp 3' UTR. The latter contains a polyadenylation signal and ends in a polyA tract. The ORFs are separated by a short intergenic region that contains at least two stop codons [10]. The bicistronic L1 mRNA is transcribed by an RNA polymerase II promoter located within the ~900 bp 5' UTR. This region also contains an antisense promoter and a third ORF (ORF0) [11, 12]. L1 integrations are typically characterised by the presence of target-site duplications (TSDs) of varying lengths [13, 14]. The first open reading frame, ORF1, encodes for a ~40 kDa protein (L1ORF1p) with RNA-binding capacity that forms flexible homotrimers and displays an RNA-binding function [15]. The monomeric L1ORF1p is composed of three distinct domains. The second ORF encodes L1ORF2p, a 150 kDa protein with endonuclease as well as reverse transcriptase activity. This protein contains an N-terminal coiled-coil domain, a central RNA recognition motif (RRM), and a C-terminal domain (CTD) [16–18]. Furthermore, the C-terminal domain contains a cysteine-rich region that is indispensable for retrotransposition activity [13].

1.4 L1 retrotransposition cycle

The L1 retrotransposition cycle represents a highly intricate process, whereby these elements can replicate and insert new copies of themselves into the host genome. The cycle commences with the transcription of a full-length L1 element from its internal promoter [19]. This consequently gives rise to the generation of a bicistronic mRNA, which encodes two distinct proteins ORF1p and ORF2p (Fig. 2). After transcription, the L1 mRNA is exported to the cytoplasm, where translation occurs. The newly synthesised ORF1p and ORF2p proteins demonstrate a marked cis preference, exhibiting a propensity to bind preferentially to their encoding mRNA, thereby forming a ribonucleoprotein (RNP) complex [20]. This cis preference is postulated to be a pivotal factor in maintaining L1 retrotransposition efficiency while limiting the mobilisation of other cellular mRNAs. Once the L1 RNP complex is formed, it must gain access to the nucleus to complete the retrotransposition process. Recent research has provided insights into the nuclear import mechanisms of L1 RNPs, suggesting roles for both passive entry during cell division and active transport through nuclear pore complexes[21]. Once in the nucleus, the L1 element initiates

target-primed reverse transcription (TPRT), a unique mechanism that couples reverse transcription with integration [22]. The process begins with the ORF2p endonuclease creating a nick in the target DNA, typically at a loosely conserved 5'-TTTT/A-3' consensus sequence. The exposed 3' hydroxyl group at the nick site serves as a primer for reverse transcription of the L1 mRNA by the ORF2p reverse transcriptase [23]. Recent structural studies have provided detailed insights into the mechanisms of L1 reverse transcriptase, revealing unique features that distinguish it from retroviral counterparts [24]. As reverse transcription proceeds, the original L1 mRNA template is degraded, and second-strand DNA synthesis occurs. The details of second-strand synthesis and the completion of integration remain areas of active research, with recent studies suggesting potential roles for host DNA repair factors in resolving the intermediates of L1 insertion [25]. The completion of this cycle results in the insertion of a new L1 copy into a new genomic location, often accompanied by target site duplications flanking the insertion. Notably, many L1 insertions are 5' truncated due to incomplete reverse transcription, resulting in insertions of variable length [26].

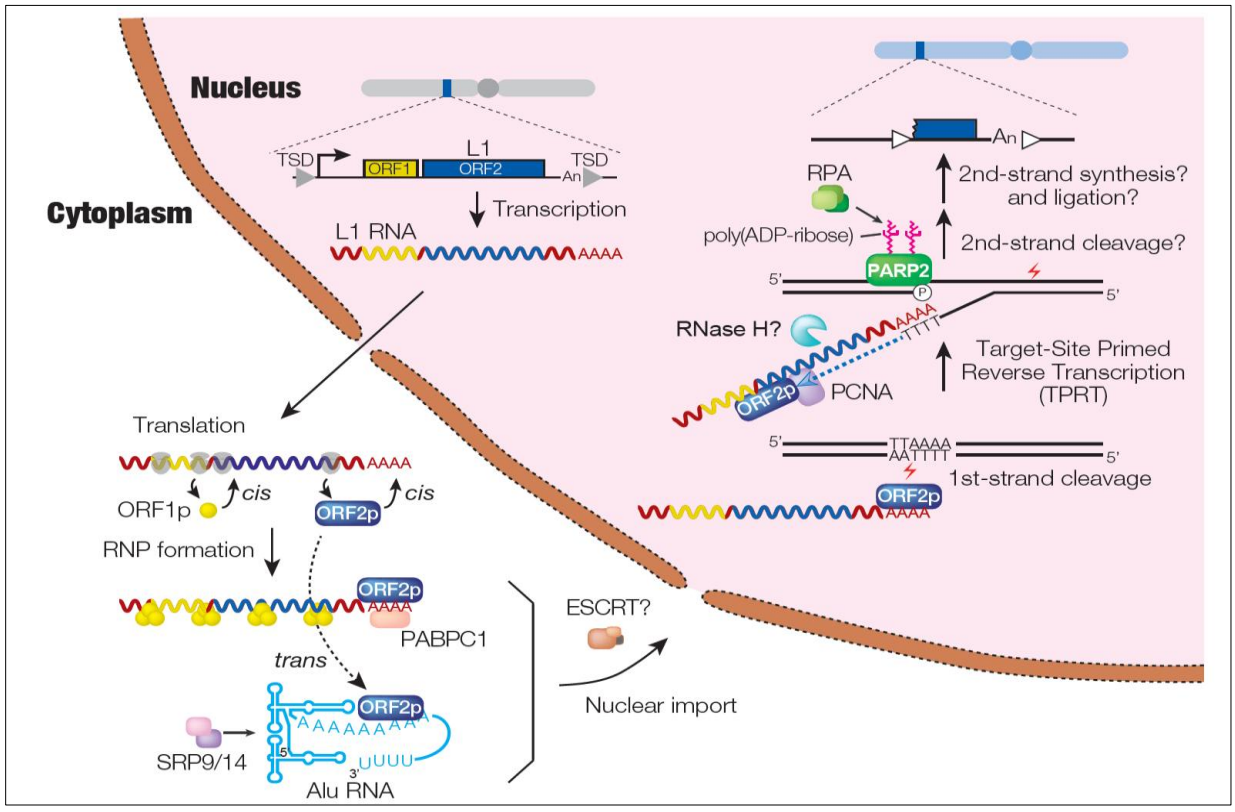


Figure 2. L1 retrotransposition cycle. (Reproduced with permission from Luqman-Fatah et al., 2022).

1.4.1 Regulation of L1 in Somatic Cells

In somatic cells, L1 retrotransposons are regulated through multiple sophisticated mechanisms that work in concert to maintain genomic stability where the entire retrotransposition cycle is subject to multiple layers of host regulation, including epigenetic silencing, RNA interference, and restriction by host factors proteins, highlighting the ongoing evolutionary arms race between L1 elements and their host genomes [27]. At the chromatin level, DNA methylation and specific histone modifications serve as primary repression mechanisms, orchestrated by key regulators including MEPC2 (methyl-CpG-binding protein 2), KAP2 (Kinase associated protein 2), sirtuins, and the NuRD (nucleosomal and remodeling deacetylase) complex [10]. Particularly noteworthy are the sirtuin genes, SIRT6 (sirtuin 6) and SIRT7 (sirtuin 7), which suppress L1 through distinct pathways. SIRT6 operates via mono-ADP ribosylation of KAP1 and interaction with MEPC2 [28], while SIRT7 functions through deacetylation of H3K18 histones and the formation of heterochromatin, which results in the association of L1 elements with the nuclear lamin A/C and their repression [29]. Post-transcriptional regulation represents another crucial layer of L1 regulation, primarily through RNA interference mechanisms. The miR-128, in conjunction with Ago protein, forms the RISC complex [10] to inhibit L1 through direct binding to the ORF2 region and by blocking transportation factors TNPO1 (Transportin-1) [30] and hnRNPA1 (Heterogeneous Nuclear Ribonucleoprotein A1) [31]. This regulation is complemented by antiviral defence mechanisms, where interferon-induced responses play a vital role. RNase L actively cleaves L1 mRNA [32], while MOV10 (Moloney leukemia virus 10) helicase forms complexes with antiviral proteins such as zinc finger antiviral protein ZAP (ZC3HAV1) which works in concert with MOV10 to block ORF1 through formation of stress granules with subsequent degradation [10]. Somatic cells also utilize host deaminase to suppress L1 activity, particularly the APOBEC3 (Apolipoprotein B mRNA Editing Catalytic Polypeptide 3) family proteins (hA3A, hA3B, hA3C, hA3F, hA3DE) [10, 33], which degrade L1 reverse transcriptase intermediates, preventing successful retrotransposition. The integration of L1 into the genome can be inhibited by proteins that are involved in the repair of single- and double-stranded DNA breaks, some studies showed that non-homologous end-joining (NHEJ) proteins such as DNA-PKCs (DNA-Dependent Protein Kinase, Catalytic Subunit) and XRCC4 (X-Ray Repair Cross Complementing 4) affecting retrotransposition [34], while DCLRE1C (DNA Cross-Link Repair 1C), LIG4 (Ligase 4), and XRCC6 contribute to L1 truncation, thus limiting the distribution of full-length active copies [35].

A recent study investigated how BRCA1 (Breast Cancer 1) ubiquitin ligase suppresses L1 retrotranspositions alongside post-replicative repair factors and homologous recombination, they found that BRCA1 employs a dual mechanism. At the replication fork, BRCA1 triggers double-stranded DNA cleavage and resection, followed by RPA (Replication Protein A) protein coating of DNA ends, ultimately creating a targeted site duplication. Additionally, BRCA1 acts in the cytoplasm to suppress both ORF2 translation and L1 RNP complex formation [36]. Cell cycle regulation adds another dimension to L1 control, with p53 exhibiting dual effects - it can increase L1 expression through promoter binding [37] while also suppressing L1 induced by p53 via activation of the piRNA (piwi-interacting RNA) signalling pathway in germ cells [38]. Furthermore, a correlation has been identified between p53 and the repressive histone mark H3K9me3 in the L1 5′UTR in zebrafish embryos [38]. In addition, p53 has been shown to inhibit L1 in cancerous and normal lung tissue from human cell cultures by binding to the 5′UTR and stimulating local deposition of repressive histone marks. The proteins p21 and p27 contribute also by binding to ORF2 protein to prevent L1 integration [39], and the MYC proto-oncogene represses L1 transcription in certain cancer cells [40].

1.5 The role of L1 in diseases and cancer

L1 retrotransposition has been recognised as a possible contributor to a wide range of human diseases, primarily through mechanisms of insertional mutagenesis and genomic instability. L1 insertion into the factor VIII gene have been identified to be a causative factor of haemophilia A [41]. Moreover, in the context of Duchenne muscular dystrophy, L1 insertion into the dystrophin gene have been detected as a potential cause of Duchenne muscular dystrophy (DMD) in several cases [42]. Moreover, there is evidence to suggest that L1 activity may be associated with several other neurological disorders including Rett syndrome, ataxia telangiectasia and autism spectrum disorders [43]. In the case of Rett syndrome, for example, the failure of MeCP2 function has been noticed to result in increased L1 mobilisation in neurons, which may contribute to the development of neurological symptoms [44]. Also, recent studies have indicated the possibility of a role for L1 in the aetiology of autoimmune diseases such as systemic lupus erythematosus (SLE) and rheumatoid arthritis. It has been demonstrated that L1-derived nucleic acids elicit innate immune responses, which may contribute to the pathogenesis of inflammatory processes [45]. Furthermore, the currently available evidence suggests that L1 activity may be involved in the aetiology of atherosclerosis and other cardiovascular diseases. The evidence suggests that L1 endonuclease

activity is associated with DNA damage in vascular smooth muscle cells, which may contribute to the development of atherosclerosis [46]. In age-related diseases, an increase in L1 activity has been observed in aged tissues and has been linked to cellular senescence, this age-related L1 activation may contribute to inflammaging and a range of age-related pathologies [47]. Numerous studies have demonstrated elevated levels of L1 expression and retrotransposition in a range of cancer types, indicating that L1 activity may play a role in the initiation, progression and genomic instability associated with cancer [48]. Somatic L1 insertions have been identified in numerous cancer-associated genes, with the potential to contribute to the inactivation of tumour suppressor genes or the activation of oncogenes. For example, L1 insertions have been identified in the APC gene in colorectal cancer and the PTEN gene in endometrial cancer [49]. Furthermore, L1 activity has been showed to induce DNA double-strand breaks, contributing to chromosomal rearrangements and overall genomic instability, a hallmark of cancer [50]. L1 insertions can affect gene expression via a number of different mechanisms, involving the introduction of new promoters, enhancers, or splice sites. This can consequence in aberrant gene expression patterns in cancer cells [51]. L1 activity can also activate cellular stress responses, including the activation of inflammatory pathways and interferon responses. These processes may participate to the tumour microenvironment and cancer progression [52]. Cancer-associated global DNA hypomethylation often causes increased L1 expression. Conversely, L1 insertions can also cause local epigenetic changes, potentially affecting the expression of adjacent genes [53].

1.6 L1 ORF1p as a biomarker for cancer

L1 ORF1p expression is not expressed in normal tissues but becomes aberrantly high in multiple cancer types. For example, Rodić et al. demonstrated that ORF1p is expressed in a wide range of human cancers, including breast, colorectal, and lung carcinomas, while remaining largely undetectable in normal tissues [54]. The presence of ORF1p in tumour samples can be readily identified through immunohistochemistry, providing a practical approach for clinical applications. The utility of ORF1p as a cancer biomarker extends beyond mere detection. Recent research has demonstrated a correlation between ORF1p expression levels and cancer progression, indicating its potential as a prognostic indicator. For example, a study by Ardeljan et al. revealed that higher ORF1p expression in estrogen receptor-positive breast cancers was associated with worse overall survival, suggesting its value in predicting patient outcomes [55]. This indicates that ORF1p could potentially serve as a biomarker for real-time monitoring of cancer progression and therapeutic

response [56]. Despite its promising potential, challenges remain in fully establishing ORF1p as a clinical biomarker. These include the need for standardisation of detection methods, determination of clinically relevant thresholds, and large-scale validation studies. Additionally, further investigation is required to determine the specificity of ORF1p for cancer versus other conditions associated with retrotransposon activation.

1.7 L1 ORF1p and tumour suppressor proteins

TP53 is the most frequently mutated gene in human cancers It guides transcriptional programmes specifying apoptosis, cell cycle arrest, and cellular senescence. Accordingly, TP53, often titled the "the genome guard", perform a fundamental role in preserving genomic stability and stopping cancer development. It is hence widely recognised that p53-mediated tumour suppression is due to the transactivation of effector genes in response to external stimuli [57, 58]. Recent studies have suggested some possible mechanisms through which TP53 and L1 elements might interact. However, the exact way in which p53 regulates L1 is still unknown. Wylie et al. demonstrated that p53 represses L1 retrotransposition, thereby proposing that loss of p53 role may contribute to increased L1 activity in cancer cells [38]. A correlation has been observed between TP53 mutations and elevated L1 ORF1p expression in a range of cancer types. For example, Rodić et al. observed a strong correlation between TP53 mutations and elevated ORF1p levels in high-grade serous ovarian carcinomas [54]. Similarly, Ardeljan et al. found an association between TP53 mutations and high ORF1p expression in colorectal cancers, demonstrating that loss of p53 function may have role in L1 activation in these tumours [25]. In addition to TP53, the retinoblastoma protein (RB1) has been exhibited to repress L1 retrotransposition, Montoya-Durango et al. demonstrated that E2F-Rb family complexes regulate L1 (L1) retrotransposons through epigenetic mechanisms, involving nucleosomal histone modifications and recruitment of histone deacetylases (HDAC1 and HDAC2) to L1 promoters, maintaining heterochromatic histone marks (H3 trimethyl K9 and H4 trimethyl K20) that suppress L1 retrotransposition. They also demonstrated that in fibroblasts lacking all three Rb family proteins (Rb, p107, and p130), L1 elements show increased histone acetylation (indicative of active transcription which means L1 expression is upregulated in these Rb family knockout fibroblasts compared to wild-type [59].

1.8 Cervical cancer

Cervical cancer is a significant international health problem, with a remarkably high incidence among women of reproductive age. It progresses in the cervix, the lower part of the uterus that links to the vagina and is often preceeded by precancerous transformations, known as cervical intraepithelial neoplasia (CIN). CIN is categorized into three grades (CIN 1, 2, and 3) depending on the extent of abnormal cell growth, CIN I is a low-grade squamous lesion (LSIL) caused by low-risk HPV types that often go away on their own in more than three-quarters of cases. CIN II and CIN III are high-grade squamous lesions (HSILs) that have a higher risk to develop to a cervical cancer. This distinction highlights the clinical importance of accurate diagnosis and the need for prompt action in cases of CIN I [60]. The World Health Organization (WHO) estimates that cervical cancer is the fourth highest common cancer in women globally, with about 604,000 new cases and 342,000 deaths stated in 2020 [61]. In developed countries, the incidence of cervical cancer has reduced significantly because of the performance of effective screening programmes. Conversely, it remains a major cause of cancer-related mortality in low- and middle-income countries. The primary cause of cervical cancer is persistent infection with high-risk types of human papillomavirus (HPV), particularly HPV 16 and 18, which comprise approximately 70% of cases [62]. Other risk factors include smoking, long-term oral contraceptive use, multiple sexual partners, and immunosuppression. The diagnosis of cervical cancer and/or CIN typically necessitates a combination of screening tests. The Papanicolaou (Pap) smear test remains a fundamental component of cervical cancer screening, frequently supplemented by HPV DNA testing for a more comprehensive risk assessment [63]. The diagnosis of CIN is primarily depended on the histological assessment of cervical biopsy samples. Though, interobserver alterability and subjectivity in the explanation of histopathological characteristics of CIN lesions can impact diagnostic precision and reproducibility [64–67]. Therefore, there is an extending need for objective and trustworthy biomarkers that can support traditional diagnostic approaches and improve the accuracy of CIN diagnosis. The most commonly utilised immunohistochemical markers for CIN are Ki67 and p16 [68]. Ki67 is a marker of cellular proliferation translated during the cell cycle and ribosomal RNA transcription [69]. In the context of cervical lesions, a high Ki67 labelling index (with more than 30% of the cells exhibiting nuclear immunopositivity) is linked with more aggressive or advanced lesions, such as high-grade CIN or cervical cancer [70]. P16, also identified as p16INK4a, is a tumour suppressor protein that acts as a negative governor of the

cell cycle, preventing the phosphorylation of retinoblastoma (Rb) protein and inhibiting cell proliferation. In high-risk HPV-associated lesions, the viral oncoproteins E6 and E7 interrupt the normal role of p53 and Rb, respectively, leading to high p16 expression as a cellular response to this dysregulation [71]. Consequently, cytoplasmic p16 overexpression has been recognized as an alternative marker for high-risk HPV infection and a reliable indicator of HPV-associated neoplasia [72]. Although these biomarkers have shown significant efficacy in diagnosing high-grade lesions, they have been less consistent and reliable for low-grade lesions, particularly CIN1. Mills et al. have published that p16 expression in CIN1 can be focal and weak, which may result in misinterpretation and false-negative results [73]. Likewise, the Ki-67 staining patterns examined in CIN1 frequently overlap with those seen in benign reactive changes, which can complicate the diagnostic process. A study by Reuschenbach et al. showed that the sensitivity of p16/Ki-67 dual staining for CIN1 was significantly lower than for CIN2+ lesions. This finding highlights the gaps in utilising these markers for low-grade dysplasia [74]. Furthermore, Wentzensen et al. observed that the specificity of p16/Ki-67 dual staining reduced considerably when employed for the identification of CIN1, which could result in overdiagnosis and unnecessary interventions [75]. These uncertainties show up the necessity for caution when utilising Ki-67 and p16 IHC for CIN1 diagnosis. Furthermore, they emphasise the importance of investigating new reliable markers to classify CIN1 cases and avoid low-grade cervical lesions misdiagnosis.

1.9 Aims of the study

The currently applied molecular diagnostic tools in cancer management rely on exome sequencing of predefined cancer driver gene panel/s. However, it is also known that a much greater segment of the genome is intronic region. Disease-causing cancer driver mutations can also be found in these large regions. For example, L1 retrotransposon integrations may create intronic driver mutations. Such L1 integration events cannot be detected by the exome sequencing procedures commonly used to date, and as a result, their significance is presumably currently underestimated. Consistent with this, recently several driver mutations were detected in different tumour types caused by new somatic L1 integration events. The concept of our research is to characterize the role of L1 retrotransposition in carcinogenesis by studying the expression pattern of L1 ORF1p in various malignant tissues and to try to find an explanation for why some cancers can worsen during treatment. Accordingly, the objective of this study is to investigate the expression of L1 ORF1p in normal and malignant tissues and to explore its potential as a biomarker for cancer progression and response to therapy. Moreover, we aim to further advance the understanding of L1 retrotransposon expression through the elucidation of the relationship between L1 ORF1p and critical tumour suppressors, particularly TP53, in tumour bank specimens. Our analysis facilitates the assessment of the correlation between ORF1p expression levels and TP53 status. This analysis is complemented by *in vitro* studies, confirming the association between L1 ORF1p and TP53. Furthermore, the objective is to evaluate the potential of L1 ORF1p as a diagnostic biomarker for cervical intraepithelial neoplasia (CIN). A systematic examination of L1 ORF1p expression patterns is conducted across a spectrum of cervical tissue samples, ranging from normal to preneoplastic and neoplastic stages, in order to assess whether L1 ORF1p levels could serve as a reliable indicator of disease progression. The progressive analysis of L1 ORF1p expression through different stages of cervical neoplasia may reveal distinct expression patterns that could aid in more accurate diagnosis and staging of CIN.

2. MATERIALS AND METHODS

2.1Tissue specimens and ethics statement

The archived pathological samples, collected for diagnostic determinations, were employed in the construction of a tissue microarray (TMA). The study was conducted on 590 samples derived from 21 distinct tumour types (Fig. 2), in addition to CIN I (n=20), CIN II (n=46), CIN III (n=14), cervical cancer (n=32), non-dysplastic cervical tissue samples (n=31) and normal tissues from various organs (n=36). The control group included 31 cases of non-dysplastic cervical tissue taken from patients who submitted to total hysterectomy for the treatment of uterine leiomyomas and uterine prolapse. Subsequently, the specimens were categorized into the following categories: normal (n = 24) and atrophic epithelium (n = 7). Subsequently, cores were selected from each case and their representativeness was evaluated following H&E staining. In the case of tumour specimens, the following parameters were investigated: tumour grade, primary tumour size, lymph node status and metastasis status (pTNM). All patient data acquired in this study were anonymised and did not affect the diagnosis or treatment plan. The age range of the patient cohort was 22 to 83 years, with a mean age of 52 years. The patients were selected from the period between 2018 and 2022 at the Clinical Centre of the Albert Szent-Györgyi Medical School at the University of Szeged, Hungary. The research protocol was checked and approved by the Institutional Committee of Science and Research Ethics of the Medical Research Council, Budapest, Hungary (reference number: BM/3049/2023).

2.2 H&E and Immunohistochemistry staining

TMA slides were incubated with three changes of xylene for paraffin cleaning, followed by rehydration through two changes each of 100%, 95%, 70%, and 30% ethanol. Once the TMA slide was removed and rehydrated, the process was repeated. Subsequently, the sections were stained in Meyer's hematoxylin solution and eosin, followed by dehydration through two changes each of 95% ethanol, 100% ethanol, and xylene. Finally, they were mounted with a xylene-based mounting medium. Immunohistochemical staining of the selected specimens was conducted on the Bond Max Autostainer (Leica Biosystems, Wetzlar, Germany) with the Bond Polymer Detection System (Vision BioSystems, Newcastle upon Tyne, UK). The formalin-fixed, paraffin-embedded tissue sections were deparaffinised, and the endogenous peroxidase was inactivated. For each antibody, antigen retrieval was conducted using either the Bond Epitope Retrieval Solution 1 or the Bond

Epitope Retrieval Solution 2 at 99–100°C for of 20–30 minutes. Subsequently, the sections were incubated with the primary antibody for 20 minutes, followed by a post-primary step for 30 minutes and then the polymer for a further 30 minutes. Subsequently, the slides were incubated with a DAB chromogen kit and counterstained with Mayer's haematoxylin. Following this, the slides were mounted and scanned with a Pannoramic Digital Slide Scanner (3D Histech). Monoclonal mouse anti-p16 antibody (clone MX007, MAD-000690QD-12, Master Diagnostica S.L, Granada, Spain) (1:200), monoclonal anti-Ki67 (clone SP6, #10080, Histopathology Ltd., Pécs, Hungary) (1:100), monoclonal anti-TP53 (clone DO-7, MA5-12557, Thermo Fisher Scientific, USA) (1:200) and monoclonal anti-L1 ORF1p antibody (clone 4H1, MABC1152, Millipore Darmstadt, Germany) (1:1000) were used. Negative controls were performed on all slides using an equivalent concentration of a subclass-matched IgG1K. The IHC sections were evaluated by 2 experienced pathologists independently to guarantee interobserver agreement. The observers were blinded to any clinical parameter.

2.3 Staining Evaluation

The intensity of L1 ORF1p staining across all tissue samples of non-cervical origin were evaluated by an experienced pathologist using a semiquantitative approach. The staining intensity was evaluated on a scale from 0 to 3, with 0 indicating no staining (negative), 1 representing weak staining, 2 moderate staining, and 3 strong staining. The immunohistochemical assessment of TP53 is a complex analytical process that extends far beyond a simple binary positive/negative interpretation. TP53 IHC staining can be scored by analysis of staining intensity, which ranges from negative through weak, moderate, to strong (0 to 3+), or by quantification of the proportion of positively stained tumour cells, usually as a percentage ranging from 0 to 100%[76]. It is also recommended to check the pattern and extent of staining across the tissue section, which provide particularly valuable diagnostic information. Three characteristic patterns have been identified as diagnostic signatures: the wild-type pattern, showing scattered weak to moderate positivity in less than 30% of cells; the null pattern, characterised by complete absence of staining, often indicating nonsense mutations; and the overexpress pattern, marked by strong, diffuse nuclear positivity in over 60% of cells, frequently associated with missense mutations [77]. As missense mutations in TP53 are significantly more prevalent in patients than nonsense mutations, the overexpress of TP53 in malignant tissues is being scored. The intensity of TP53 overexpress in our cohort was categorised on a scale from 0 (normal expression) to 3 (strong overexpress). In order to evaluate

the ORF1p immunostaining results for a sample of cervical origin, a modified version of the immunoreactive scale (IRS) by Remmele and Stegner was performed [78]. The samples were scored according to both the intensity of the staining and the extent of epithelial involvement. The intensity of cytoplasmic staining was evaluated using a semi-quantitative scale. The scoring system employed was as follows: 0 (absent), 1 (involving the basal one-third), 2 (involving the basal two-thirds), and 3 (involving the entire epithelial thickness). To ascertain the ORF1p immunoreaction score, the intensity and extent scores were multiplied. If less than 10% of the epithelial cells exhibited weak reactivity, the ORF1p staining was considered to be negative. In order to evaluate the value of Ki67 immunoreactivity, the nuclei of 200 epithelial cells, distributed across the entire epithelial layer, were explored in each specimen. The Ki67 index was defined as the percentage of cells that exhibited positive staining for Ki67. The grades of 1, 2, and 3 were assigned for the purpose of categorizing the extent of expression. The Ki67 index was classified as Grade 1, 2, or 3 when the percentage was below 5%, between 5% and 30%, or above 30%, respectively. The extent of the p16 immunoreaction was determined by calculating the percentage of p16-positive cells. In accordance with the semi-quantitative scale, a designation of 0 was assigned when the percentage of positive cells was below 1%. Cases demonstrating clustered positive cells with a percentage of positive cells ranging from 1% to 5% and from 5% to 25%, respectively, were assigned grades 1 and 2, correspondingly.

2.4 Plasmid construction

To study the effects of P53 knockdown on L1 retrotransposition in cell culture, we modified a construct we have already developed [79]. In this the neomycin selective marker gene was cloned to the A side (HADHA) of the human hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex alpha and beta bidirectional promoter. The neomycin gene is disrupted with a short intron from which an artificial microRNA (amiR) is expressed to silence the TP53 gene. On the B side (HADHB) of the bidirectional promoter we cloned an ORFeus-type L1 reporter element[80]. These transcriptional units were cloned between the two inverted terminal repeats (ITRs) of the Sleeping Beauty transposon.

2.5 Cell culture, transfection, fluorescence-activated cell sorting (FACS) analysis and paclitaxel treatment

The impact of TP53 expression on L1 retrotransposition was investigated in HepG2, HT-1080 (with normal TP53 function) and Saos-2 (with TP53 deletion) cell lines. HepG2 cells were cultured in Dulbecco's modified Eagle's medium (DMEM), supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (P-S). The cells were seeded at a density of 3.0×10^5 per well in a six-well plate and cultured in a humidified atmosphere with 5% CO_2 at 37 °C. The cells were transfected with 500 ng of the L1 ORFeus reporter construct in combination with 50 ng of the hyperactive SB100 transposase helper plasmid, using FuGENE® HD transfection reagent (Promega) in accordance with the manufacturer's instructions. To establish stable cell lines, transfected cells were selected with neomycin 500 µg/mL (G418). The rate of L1 retrotransposition was measured with FACS. For these 20,000 cells per sample were analysed once a week for a period of 2-3 months following transfection. This was conducted using forward scatter versus green fluorescence plots (FACSCalibur instrument; BD Biosciences, Sparks, MD, USA). The gate for EGFP-positive cells was determined by analysing cells transfected with an EGFP-negative expression plasmid (pCEP-Puro). To test the effect on cancer therapy on L1-ORF1p expression, cells were treated with 5 nM final concentration of paclitaxel in the culture medium.

2.6 Data analysis

The data was processed and analysed using the GraphPad Prism software (version 9.4.1 for Windows, GraphPad Software), the R statistical environment and the Python programming environment. Comparisons were made between the entities in question using Pearson's chi-squared test, the chi-squared independent test and Fisher's exact test. In all cases, an alpha level of 0.05 was used to determine whether the observed associations were statistically significant.

3. RESULTS

3.1 L1 ORF1 PROTEIN EXPRESSION IN SOMATIC TISSUES

3.1.1 Sample selection and analysis

Following tissue processing and immunohistochemistry staining, samples with missing sample cores or lacking clinicopathological data were excluded from the study. The analysis comprised 590 samples derived from 21 distinct tumour types (Fig. 3), in addition to CIN I (n=20), CIN II (n=46), CIN III (n=14), cervical cancer (n=32), non-dysplastic cervical tissue samples (n=31), and normal tissues from various organs (n=36). All specimens were formalin-fixed paraffin-embedded (FFPE) and tested for L1 ORF1p immunopositivity.

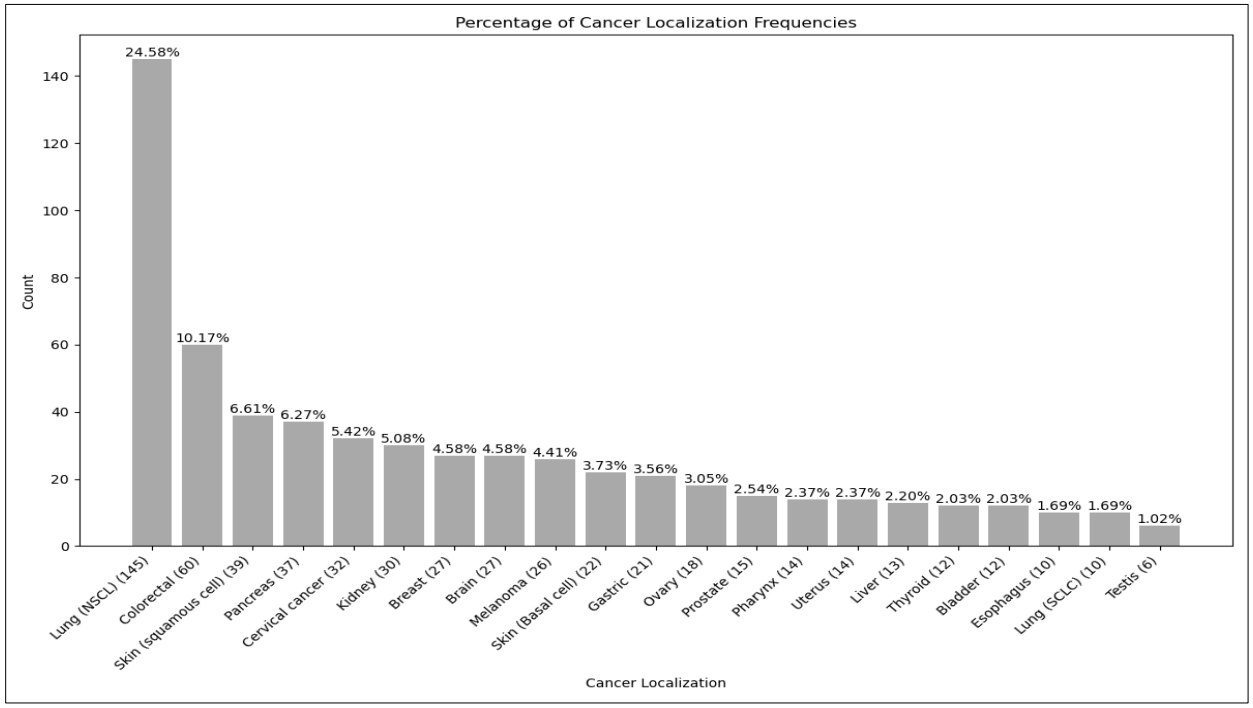


Figure 3. The frequencies and types of tumour specimens included in this study.

3.1.2 L1 ORF1p is not present in normal somatic tissues but is expressed in malignant tissues

In order to ascertain the expression pattern of the L1 ORF1p in normal tissues, we employed IHC to survey L1 ORF1p expression in mature somatic tissues from two male and two female autopsy cases on 36 different organs. The staining conditions were validated using adult testis, which is reported to show L1 protein expression [81], and which exhibited low levels of immunolabeling in maturing spermatogonia. The results obtained demonstrated low levels of expression of L1 ORF1p in striated muscles, including skeletal muscles, myocardium, and oesophageal striated muscle fibres. Additionally, intermediate to high ORF1p expression was observed in the stratum

basale of the epidermis of several normal skin specimens. Conversely, the majority of other normal mature somatic tissues exhibited an absence of expression of L1 ORF1p (Fig. 4). The ORF1 protein is predominantly located in the cytoplasm, where it binds to L1 mRNA to form stable ribonucleoprotein (RNP) complexes. This finding is corroborated by the results presented in (Fig. 5), which demonstrate that L1 ORF1p is frequently detected in the cytoplasm (and on occasion in the nucleus) of a diverse range of human cancers, with no discernible presence in mature human somatic tissues. Furthermore, approximately 355 of the 590 cases (57%) exhibited a notable degree of intermediate and high ORF1p immunoreactivity, which is a considerably higher proportion than the 161 cases (27%) that demonstrated no ORF1p expression.

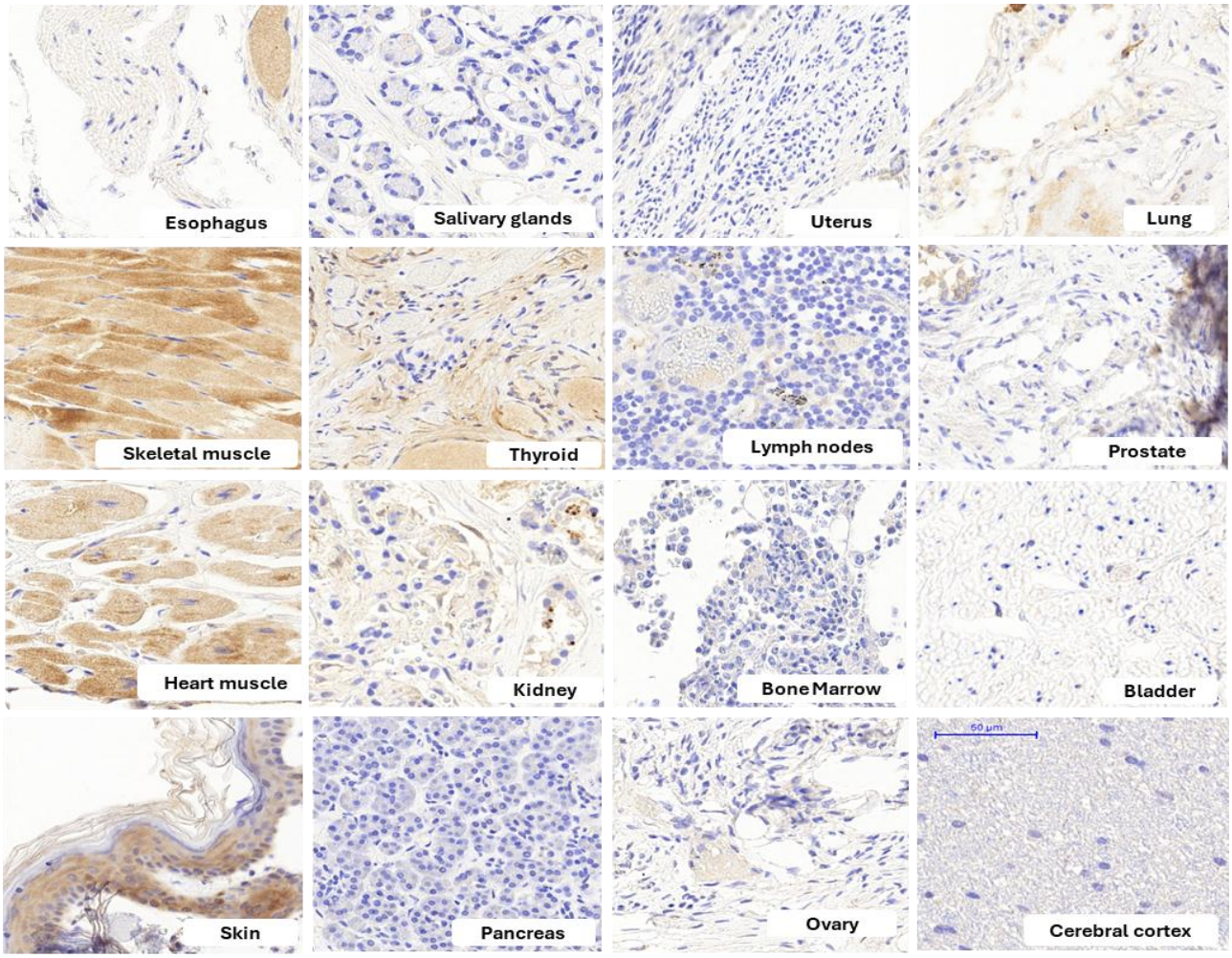


Figure 4. L1 ORF1p is not expressed in mature human somatic tissues. L1 ORF1p immunoreactivity shown as intracytoplasmic brown pigment, scale bar = 50 um.

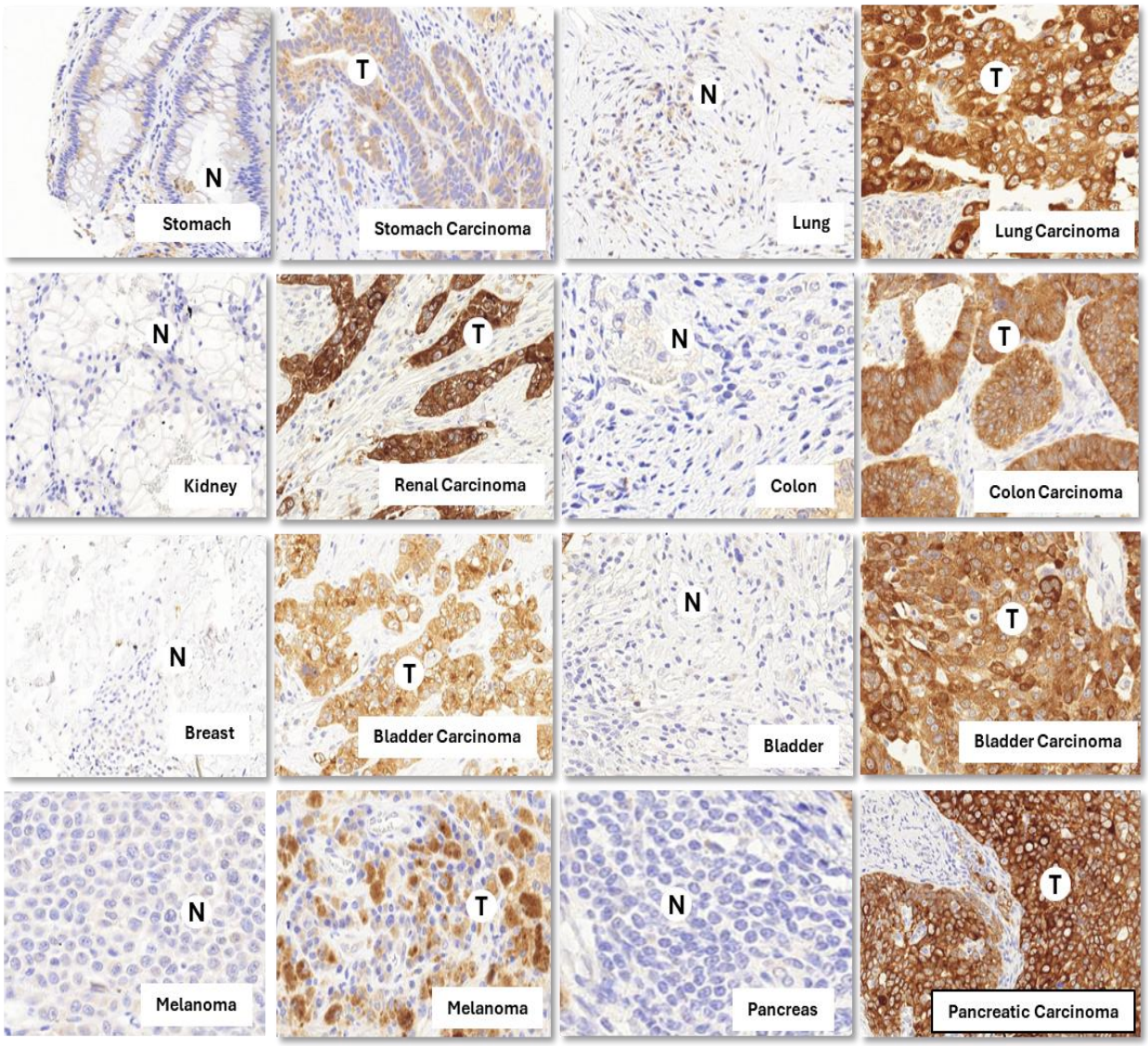


Figure 5. L1 ORF1p is undetectable in mature human somatic tissues and commonly expressed in a wide range of human cancers.. N= normal, T=tumour, scale bar is 50 um.

The immunohistochemical staining of FFPE specimens with ORF1p polyclonal antibody demonstrated a significant variation in ORF1p expression based on the primary site of origin across different cancer localisations. It is notable that a considerable proportion of cancers exhibited a marked elevation in ORF1p expression. This included skin basal cell carcinoma (100%), cervical cancer (83.87%), oesophageal cancer (70%), and non-small cell lung cancer (59%) (Fig. 6).

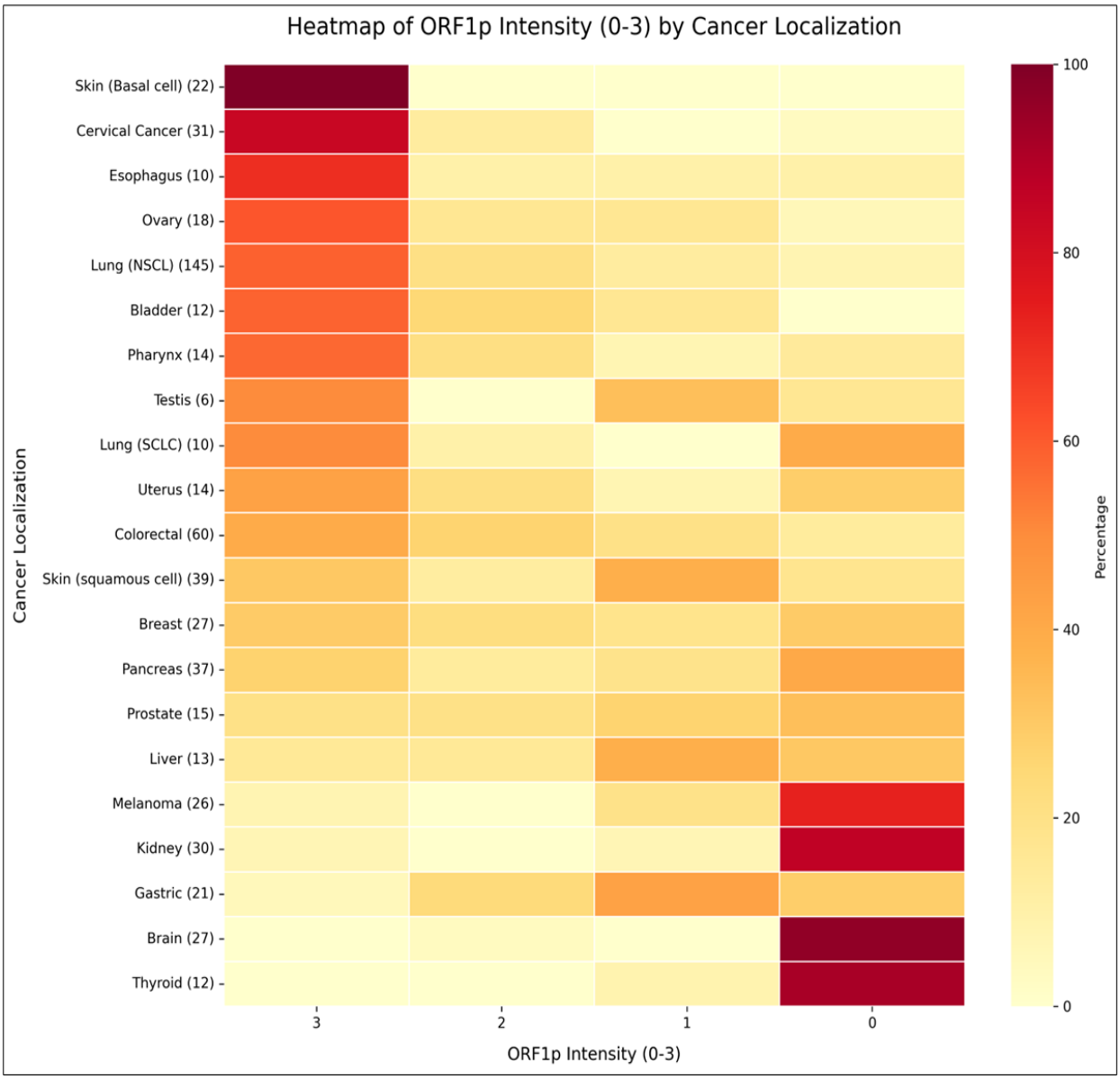


Figure 6. The level of ORF1p (0-3) expression differs between various types of cancer The numbers in parentheses indicate the sample size for each cancer type. Darker shades indicate higher percentages N = 590.

3.1.3 L1 ORF1p expression correlates with cancer progression

The subsequent analysis examined the expression pattern of L1 ORF1p across different clinical stages and progression of cancer. The findings revealed that the L1 ORF1p expression was elevated in high-grade tumours and advanced stages in comparison to low-grade and non-advanced tumours. Furthermore, a notable observation was the highest number of Grade 3 cases (n=125) with an ORF1p intensity of 3 (Fig. 7C). This trend is consistent across all grades, with the number

of cases increasing in tandem with both grade and ORF1p intensity. In contrast, cases with low ORF1p intensity are more frequent in lower-grade tumours (Fig. 7C). Similarly, the highest frequency of cases (n=65) is observed in Stage 4 with an ORF1p intensity of 3 (Fig. 7A). Additionally, it was observed that Stage 1 exhibited a relatively uniform distribution across all ORF1p intensities, whereas Stages 2-4 demonstrated greater variability (Fig. 7A). The non-significant p-value (p=0.363) indicates that the observed differences in stage frequencies across ORF1p intensities do not reach statistical significance, suggesting a potential association between advanced disease and elevated ORF1p expression.

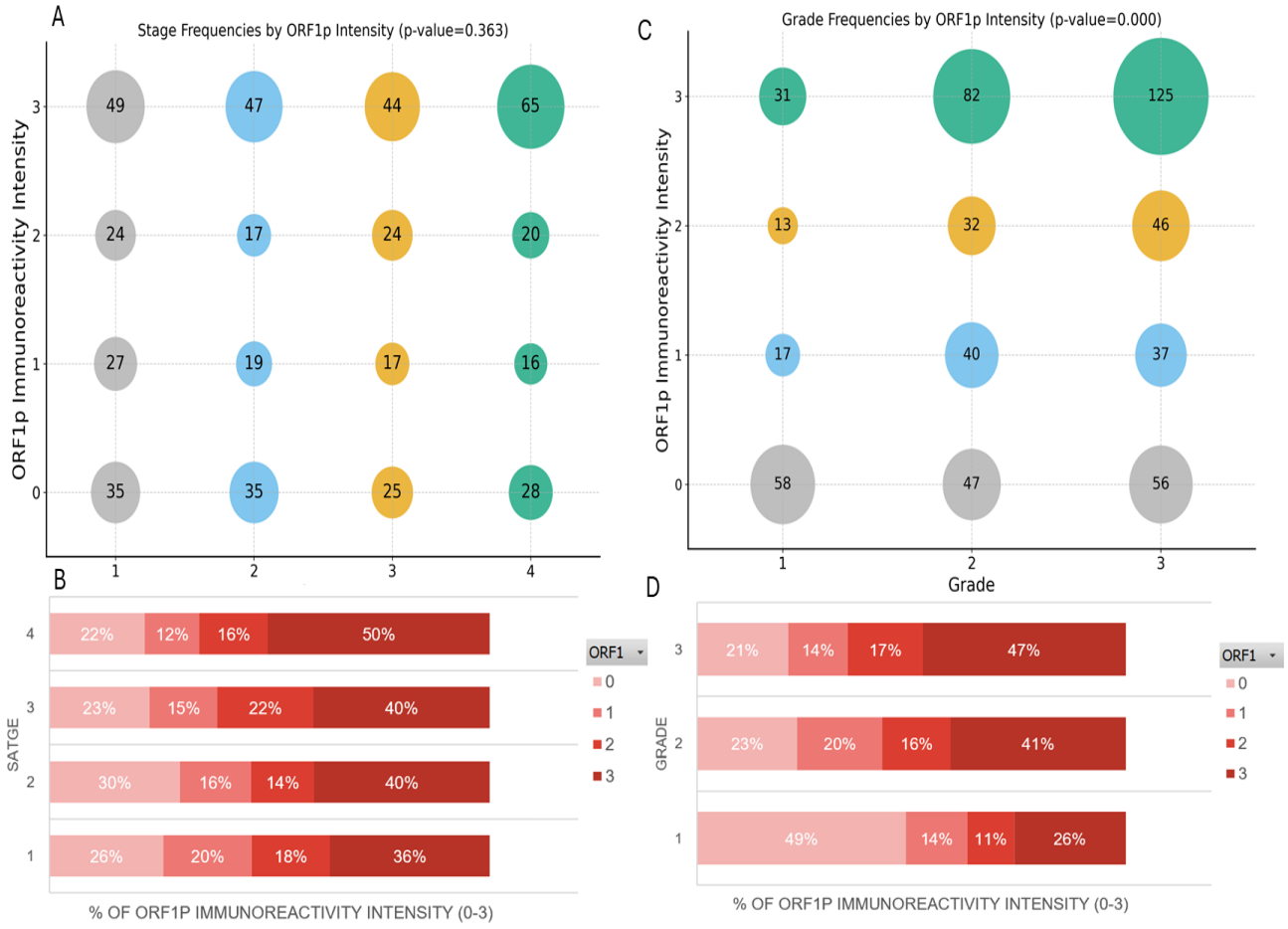
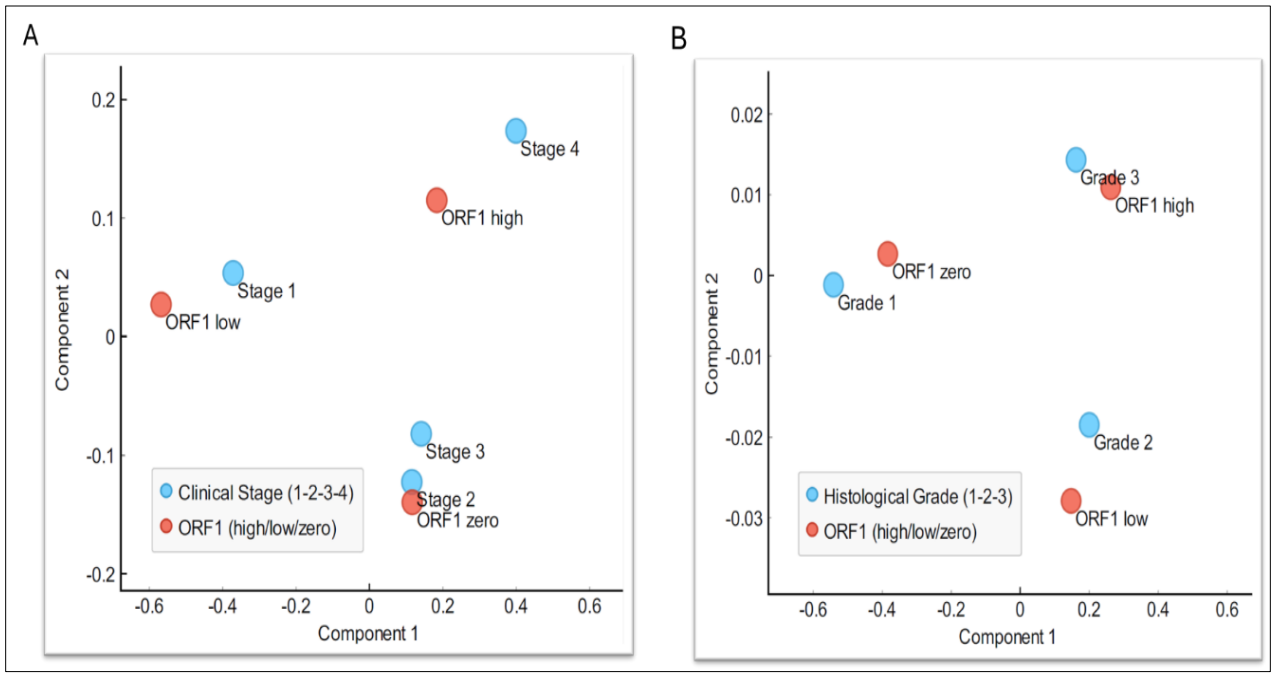


Figure 7. L1 ORF1p expression correlates with cancer progression. (A & C) Distribution of cancer cases by clinical stages (1-4), histological grades (1-3) and L1 ORF1p immunoreactivity intensity (0-3). The size and number within each bubble indicate the number of individuals exhibiting a specific immunoreaction score within each tissue sample group. (B & D) Frequencies distribution of the percentage immunoreactivity intensity of L1 ORF1p among clinical stages (1-4) and grades (1-3) of cancer N = 590.

To investigate the relationship between ORF1p immunoreactivity levels and both clinical stages and histological grades of cancer, we conducted a Principal Component Analysis (PCA) (Fig. 8A). This analysis revealed a clear separation of clinical stages (Stage 1-4) and histopathological grades (Grade 1-3) along with ORF1p immunoreactivity levels (low, high, zero). The analysis demonstrates that the various clinical stages are distinctly separated in the PCA space, with a clear progression from Stage 1 to Stage 4 along the Component 1 axis. The "ORF1p low" group is situated in closer proximity to Stage 1, while the "ORF1 high" group is positioned in the vicinity of Stage 4, which suggests a potential correlation between elevated ORF1p levels and more advanced clinical stages. Similarly, the "ORF1 high" group is located close to Grade 3, indicating that elevated ORF1p levels are associated with more aggressive tumour grades (Fig. 8A). Conversely, the "ORF1 zero" and "ORF1 low" groups are situated closer to Grade 1 and Grade 2, indicating that lower ORF1p levels may be associated with less severe grades. Additionally, the expression of L1 ORF1p was found to be positively correlated with higher histological grade of the tumours ($r = 0.2$, $p < 0.05$) as well as advanced pathological stages ($r = 0.11$, $p < 0.05$) in our cohort (Fig. 8C&D). This suggests a modest but statistically significant correlation, particularly between higher ORF1p (ORF1p 3) and advanced cancers (stage 4 and grade 3), as indicated by the heat map (Fig. 8B).



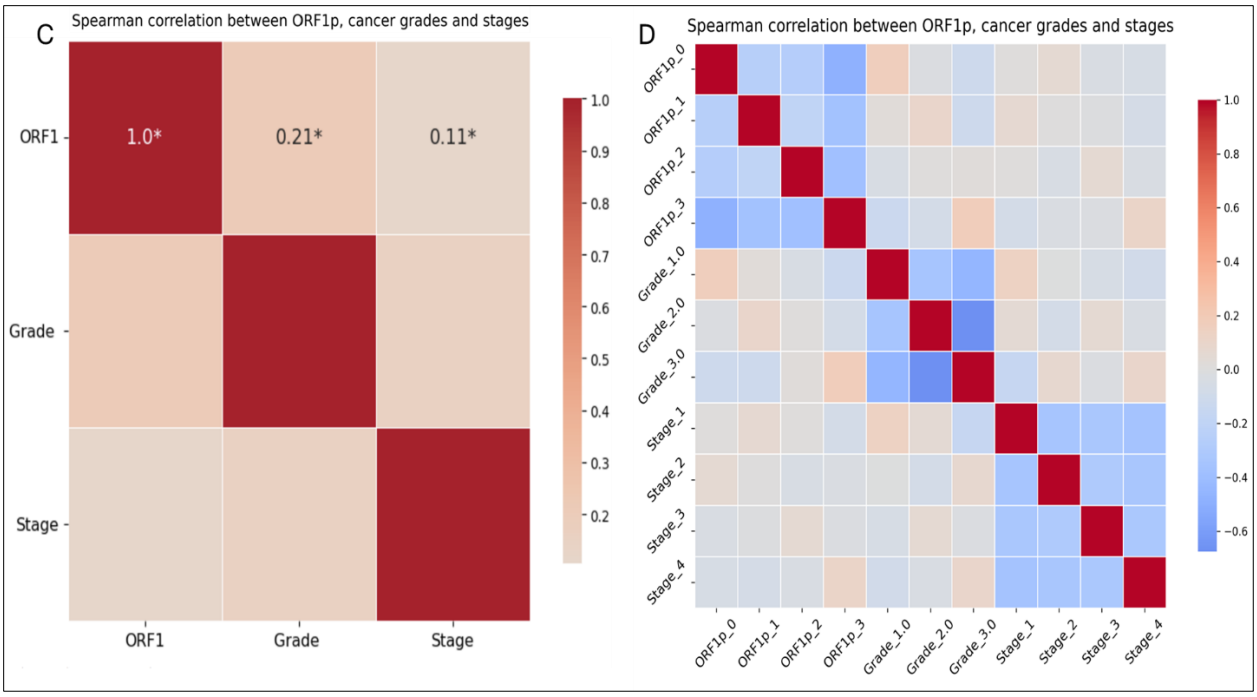


Figure 8. The level of ORF1p immunoreactivity is elevated in advanced clinical stages and higher histological grades. The points on the Principal Component Analysis (PCA) represent combinations of ORF1p levels with either clinical stages (A) or histological grades (B). (C&D) L1 ORF1p immunoreactivity is positively correlated with clinical stages and histological grades. N = 558.

3.1.4 L1 ORF1p expression shows intratumoral heterogeneity in some cancers

It is well established that tumour heterogeneity can be a significant factor in tumour progression [82]. An increase in L1 activity in undifferentiated cells and areas is associated with an advancement in tumour stage, which is a key indicator of progression. Consequently, it is essential to consider the potential for heterogeneity in certain cancers. The objective was to ascertain the role of the aforementioned types in tumourigenesis. To this end, we initially analysed L1 ORF1p expression by IHC on tissue microarray (TMA) cores in a wide range of tumour types. This revealed spatial inhomogeneity in the staining pattern in certain samples. A significant degree of cellular variability was observed within the tumours, with some cells exhibiting high ORF1p expression in the peripheral regions while others showed low expression in the central region. Additionally, unicellular heterogeneity was identified, where individual cells exhibited considerable differences in ORF1p expression levels. Of particular note was the observation of high ORF1 staining intensities at the solid component of endometrial cancers. Since only a small piece of tissue from the donor block is used for the preparation of the TMA slide, intratumoral

heterogeneity can greatly affect the evaluation of TMA. To overcome this, we evaluated whole slides from tumours which showed the most pronounced differences in ORF1p staining pattern, such as endometrial (EC), colorectal (CRC) and basal cell carcinoma of the skin (BCC). This approach aimed to provide a more comprehensive representation of the tumour heterogeneity, which may not be accurately reflected by the use of TMA slides. Our findings revealed that over half of the samples (63%) exhibited heterogeneity, manifested in a mosaic pattern (22%) or a central-peripheral distribution (41%) of high ORF1 staining intensity. This mosaic-type pattern was either unicellular (Fig. 9 B & C) or characteristic of solid (Fig. 9 D & E) in comparison to the typical ORF1p positivity observed in normal endometrial tissues (Fig. 9 A). The solid type appeared to mirror the tumour dedifferentiation, with the L1 ORF1p expression being associated with tumour progression. In endometrial cancer, the solid component represents an architectural hallmark and criterion of dedifferentiation. Higher expression of ORF1p was observed in solid components compared to glandular lower-grade tumour parts (see Figure 9, panels D and E). In certain instances of BBC, the ORF1-expressing cells exhibited pronounced invasive properties (indicated by arrows) in comparison to the negative or low-expressing tumour cells (Fig. 9 F).

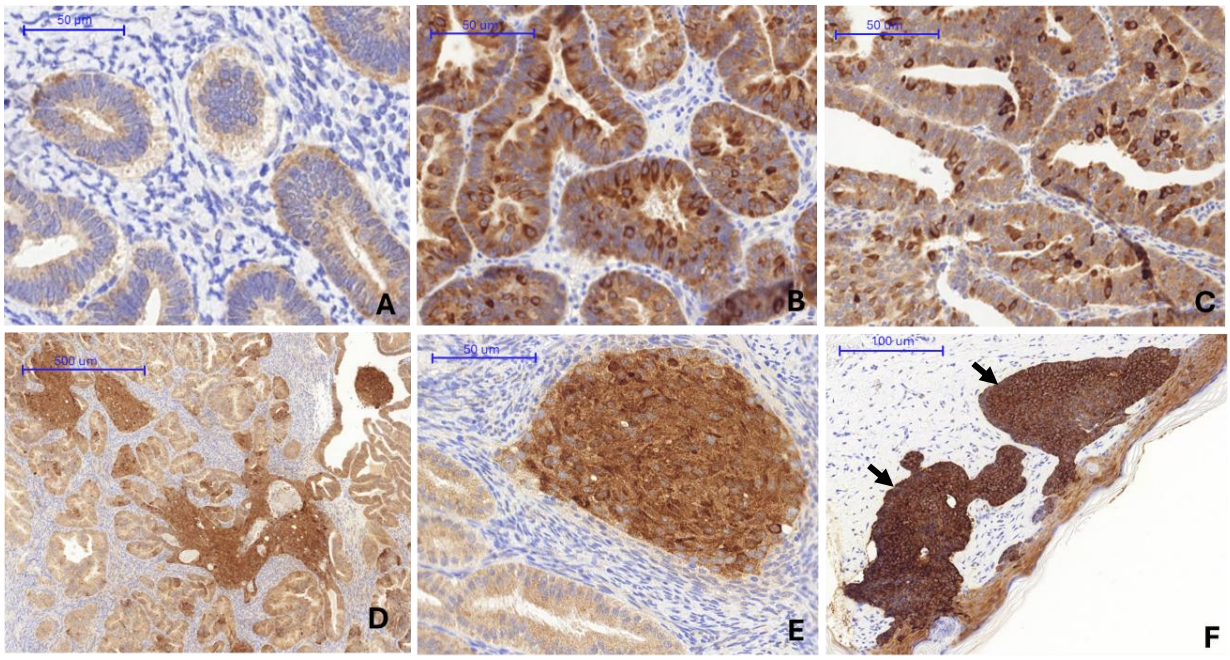


Figure 9. Intratumoral heterogeneity of L1 ORF1p expression. Some cells show high ORF1p expression in the peripheral regions, others show low expression in the central region. (A) Normal endometrial tissue with normal OFR1p immunoreactivity. (B) Endometrial carcinoma with heterogeneous ORF1p

immunoreactivity. (C) Endometrial carcinoma with mosaic unicellular ORF1p heterogeneous expression. (D) & (E) Endometrial carcinoma with mosaic heterogeneous ORF1p expression in solid vs granular tumours. (F) basal cell carcinoma, ORF1p-expressing cells show marked invasive properties.

3.2 L1 ORF1p AND TUMOUR SUPPRESSOR GENES

3.2.1 L1 ORF1p expression is increased in TP53 immunoreactive malignant tissues

To investigate the hypothesis that acquired mutations in certain tumour-suppressor genes may correlate with L1 ORF1p expression, our cohort samples were also stained with anti-human P53 antibody by IHC. We found that ORF1p expression occurred more frequently in the absence of functional p53 protein where TP53 immunoreactive area also exhibited high immunopositivity for ORF1p (Fig. 10A). Among the cases exhibiting TP53 immunoreactivity (n=267), 89.1% demonstrated ORF1 immunoreactivity, while only 10.9% exhibited ORF1 non-immunoreactivity (Fig. 10B). In contrast, among TP53 non-immunoreactive cases (n=277), only 54.9% exhibited ORF1 immunoreactivity, while 45.1% were ORF1 non-immunoreactive. The observed difference in ORF1 immunoreactivity between TP53-positive and TP53-negative groups is statistically significant (p=0.0000), indicating a potential association between TP53 and ORF1, since if TP53 is immunoreactive, it is much more likely that ORF1 also immunoreactive (Fig. 10B).

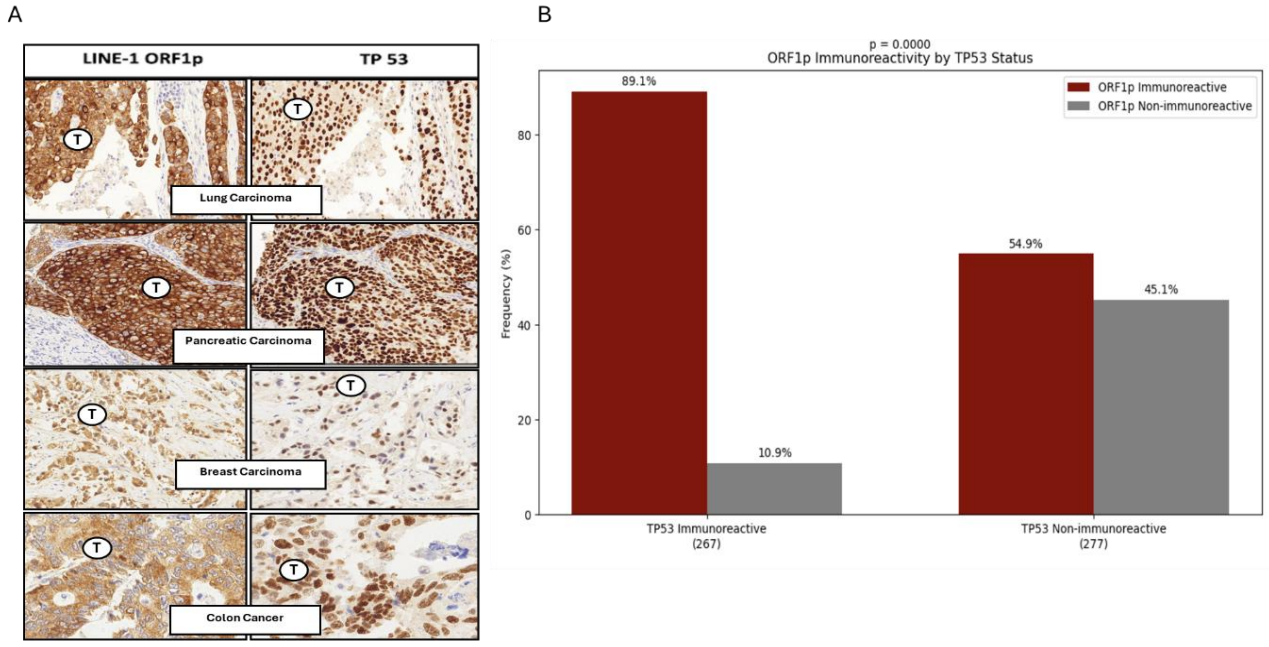


Figure 10. L1 ORF1p expression is increased in TP53 immunoreactive malignant tissues. (A) L1 ORF1p immunoreactivity is shown as intracytoplasmic brown pigment while TP53 immunoreactivity shows the

over-expression of a mutated protein as intranuclear brown pigment. (B) ORF1p immunoreactivity in TP53 immunoreactive and TP53 non-immunoreactive samples. The P-value was analysed using a Chi-square test of independence, scale bar = 50 um, T = tumour, N = 544.

Then, we further elucidated the relationship between ORF1p expression in tumours with non-functional TP53. The regression analysis was employed to compare the overall frequencies and percentages of ORF1p and TP53. A moderate positive correlation was observed between ORF1p and TP53 immunoreactivity (Spearman r = 0.46, p = 2.659e-30) (Fig. 11A). The regression line shows the overall trend of the relationship between ORF1p and TP53. As ORF1p immunoreactivity increases, TP53 tends to increase too. The data points are dispersed around a general upward trend, which suggests a moderate positive relationship between the two variables (Fig. 11A). Subsequently, we conducted a further analysis of the correlation between each intensity of ORF1p and TP53 immunoreactivity. The most notable correlation was observed between ORF1_0 and TP53_0 and ORF1_3 and TP53-3 where they demonstrated moderate positive correlation of 0.38 and 0.33, respectively (Fig. 11B).

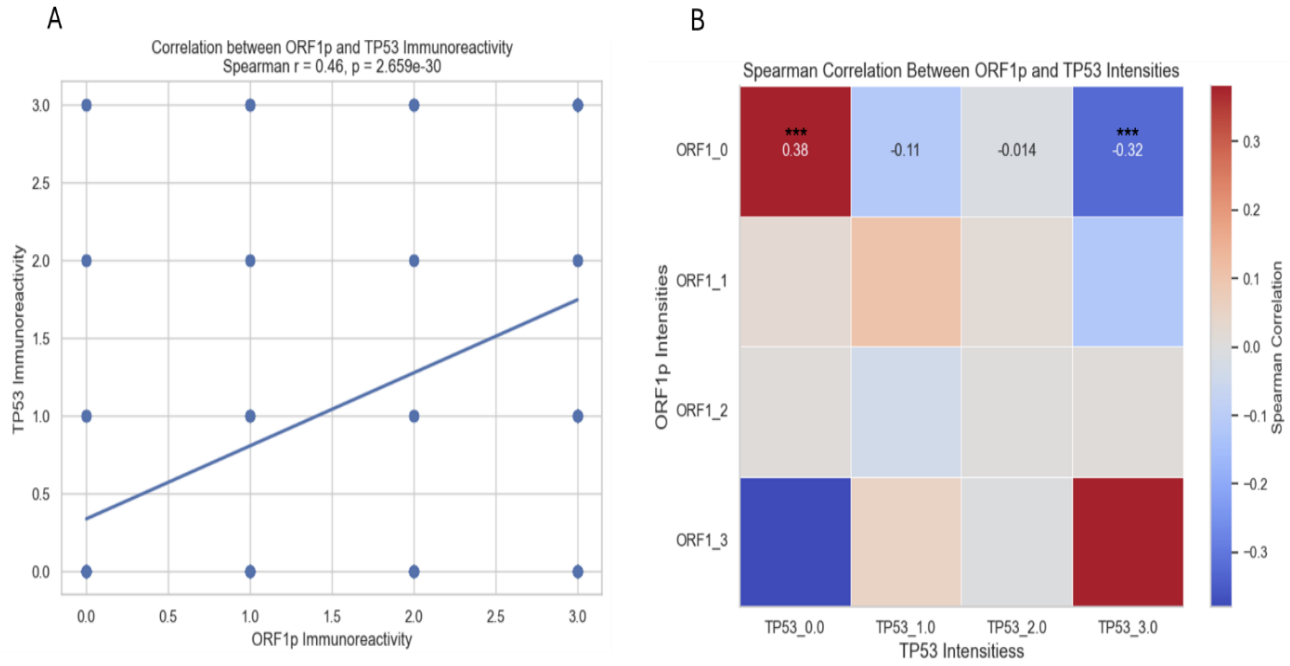


Figure 11. L1 ORF1p expression is correlated to TP53 overexpression. (A) The scatter plot shows how ORF1 and TP53 levels are correlated. (B) The heatmap shows the correlation between different levels of ORF1p and TP53. *** indicates p < 0.0001 (statistically significant) N=544.

To gain further insight into the relationship between L1 ORF1p immunoreactivity and TP53 overexpression, we conducted a comprehensive analysis of the expression patterns of both variables across a diverse range of cancers. Initially, we compared the percentage distribution of ORF1p and TP53 immunoreactivity intensities (0-3) across 21 distinct cancer localisations. Our findings revealed that a considerable number of cancer types exhibited a high prevalence of samples with strong immunoreactivity (3). In a number of cancers, including skin (BCC), colorectal, ovarian, pharynx and oesophagus cancers, as well as NSCLC and bladder cancers, both ORF1p and TP53 are highly expressed. Conversely, in other cancers, such as brain, thyroid, kidney and prostate cancers, as well as melanoma, high ORF1:0 and TP53:0 percentages are observed, but low ORF1:3 and TP53:3 percentages. The variability of ORF1p immunoreactivity across other cancer types is greater than that of TP53, and there is no distinct association between the two (Fig. 12A). Secondly, a PCA biplot was performed to confirm and provide a comprehensive overview of the linked relationships between ORF1 and P53 immunoreactivity. The results showed that skin (BCC), colorectal, oesophageal cancers, and NSCLC are most strongly linked to high ORF1p and TP53 immunoreactivity (ORF1:3 and P53:3), where they are grouped close together in the lower right part of the plot. This is close to the lines for ORF1:3 and P53:3. This shows that these cancers are positively linked to high levels of both ORF1p and TP53 immunoreactivity (see Fig. 12B).

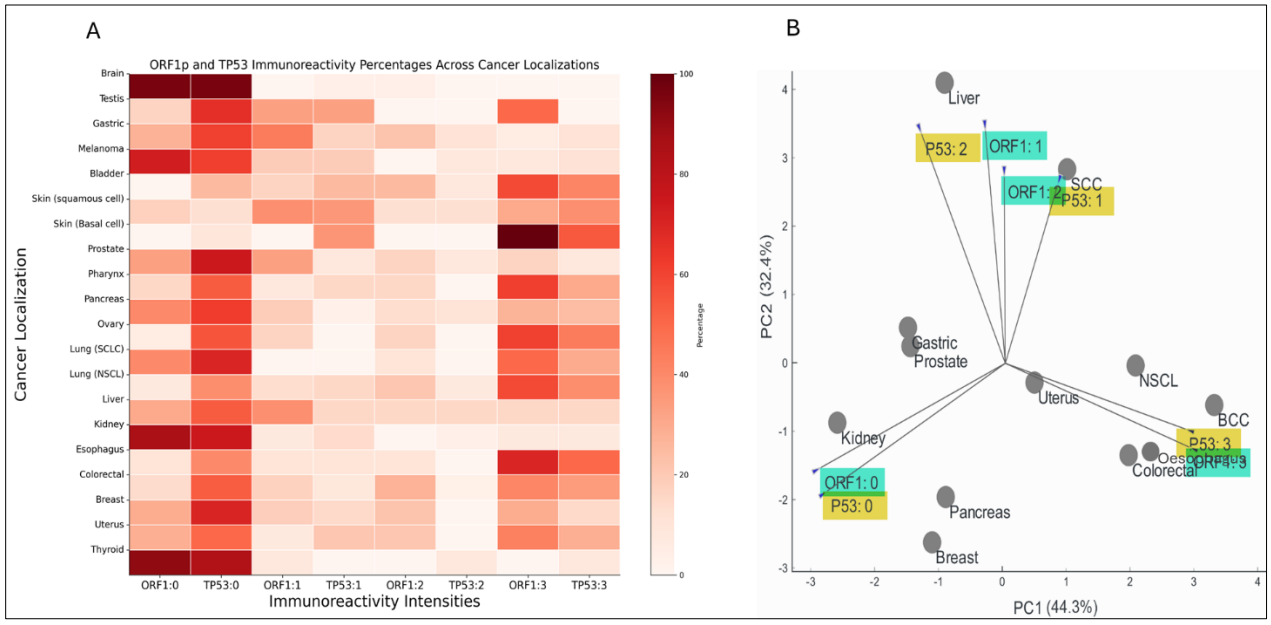


Figure 12. High ORF1p expression is commonly linked to TP53 overexpression in BCC, oesophagus, CRC and NSCLC. (A) Heatmap shows how often ORF1p and TP53 are found in different types of cancer N=558.

(B) PCA of ORF1 and P53 immunoreactivity across various cancer types. The x-axis represents PC1, accounting for 44.3% of the total variance, while the y-axis represents PC2 which shows 32.4% of the total variance. Samples that are ORF1 positive but P53 negative are not included in this analysis, so N = 134.

3.2.2 Tumors with mutated BRAC1/2, HER2, KRAS showed increased ORF1p immunoreactivity.

Given the positive correlation between ORF1p expression and TP53 overexpression, we examined the correlation of ORF1p expression with that of other tumour suppressor genes and oncogenes. Prior to the collation of personalised treatment options from the medical records of the Institute of Pathology, Faculty of Medicine SZTE, the results of the oncogene sequencing panel were collected from the patients' records and samples stained with ORF1p antibody. The results indicate that ORF1p immunoreactivity varies significantly with the mutational status of certain cancer-related genes. The strongest associations are observed with BRCA1/2, as cancers with mutated BRCA1/2 demonstrated significantly elevated ORF1p immunoreactivity (86%) compared to BRCA1/2 wild-type cancers (47%) (Fig.13 A). This strong association suggests the potential for an interaction between BRAC1/2 mutations and ORF1p expression. Also, HER2 mutated cancers exhibit high ORF1p immunoreactivity (54%) compared to HER2 wild-type cancers where 100% of cases were not expressed ORF1p (Fig.13 C). Conversely, KRAS and EGFR mutations appear to have a limited impact on ORF1p immunoreactivity (Fig.13 B & D). In general, the observed variability in the results suggests the existence of intricate interactions between ORF1p expression and various oncogenic pathways, which may have implications for the further understanding of cancer biology.

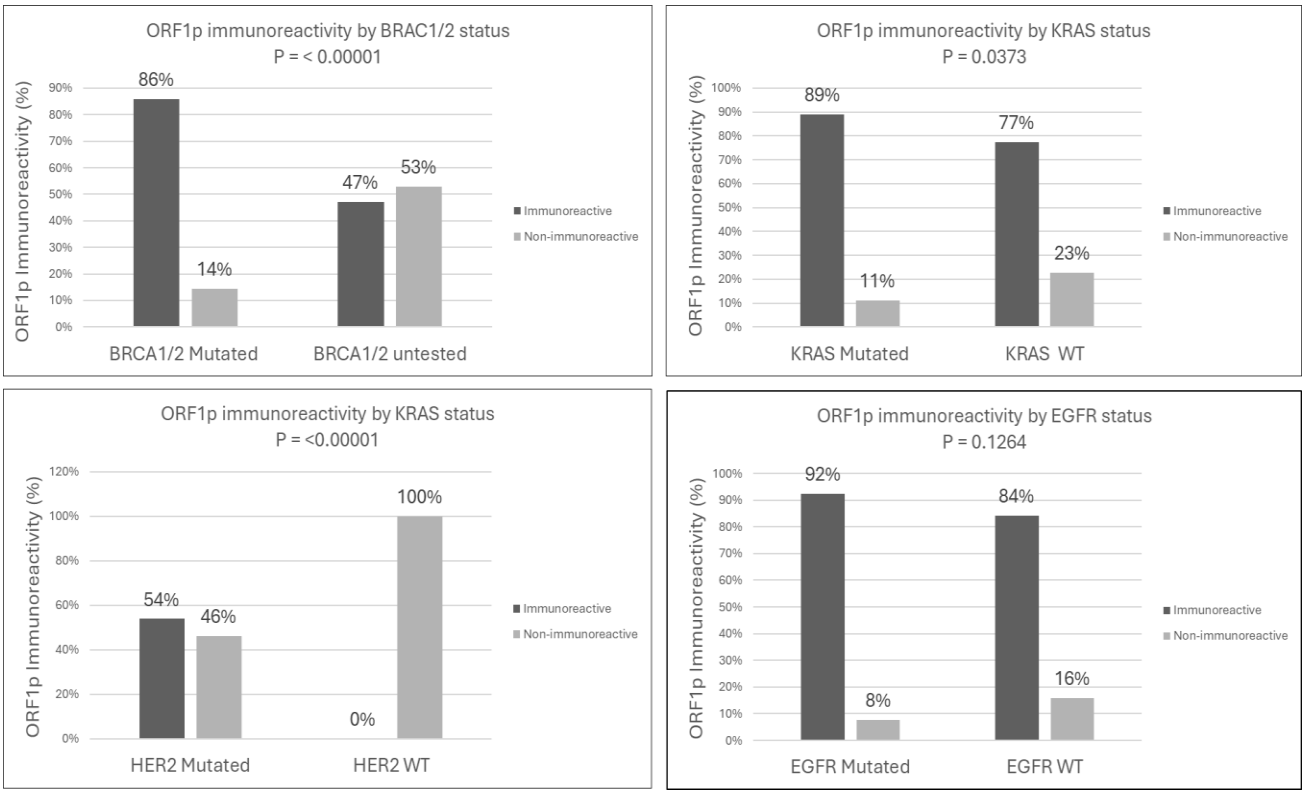
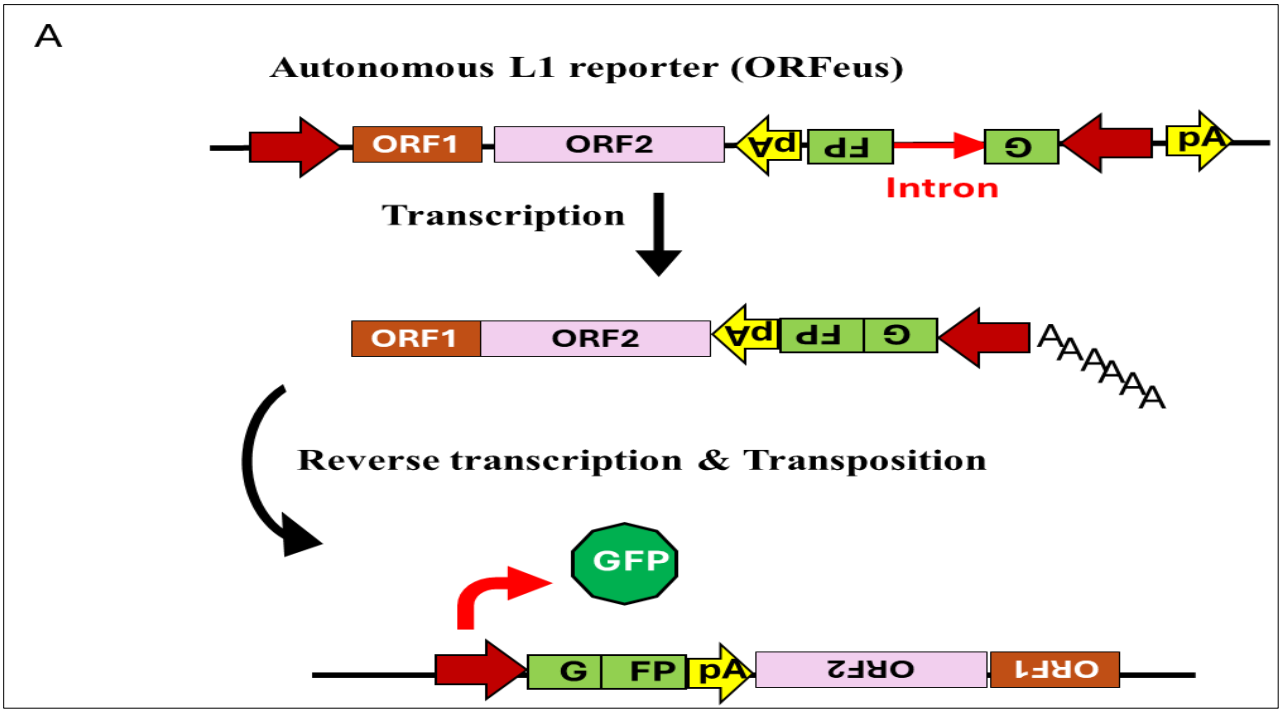


Figure 13. Tumours with mutated BRAC1/2, HER2, KRAS & EGFR (n = 24, 49, 16 & 34 respectivly) displayed increased ORF1p immunoreactivity.

3.2.3 Endogenous TP53 silencing in vitro increases the rate of L1 retrotransposition

To corroborate the findings of the tumour bank study, we developed an *in vitro* assay system to investigate the effect of endogenous TP53 gene silencing on L1 retrotransposition, using the methodology described in Kopasz et al [79]. This system allows to investigate the impact of endogenous gene silencing on L1 retrotransposition, by establishing stable mammalian cell lines expressing, in addition to an ORFeus-type synthetic human L1 reporter, a selective marker gene and an amiR from an inserted intron for the silencing of any arbitrary gene. This arrangement is cloned between the ITRs of the SB transposon and stable chromosomal gene transfer is performed by the SB100 hyperactive transposase [83]. The L1-ORFeus expression plasmid contains an EGFP cassette in the L1 3′ UTR, which is in the opposite orientation to the L1 and is interrupted by an intron. The expression of EGFP is only observed in cells subsequent to the splicing and intron removal of the L1-ORFeus transcript, reverse transcription, and integration into chromosomal DNA (Fig. 14 A &B). After retrotransposition, an EGFP signal is generated from the reporter,

which can be measured by FACS (Fig. 14 A &B). HepG2 (human hepatocellular carcinoma) and HT-1080 (human fibrosarcoma) malignant cell lines, with normal TP53 function, were transfected with either the aforementioned construct or its amiR-less version. Additionally, the Saos-2 (human osteosarcoma) cell line, which does not express p53 protein due to a homozygous deletion of its gene, was transfected with the same constructs. This cell line represents an important control model for investigating the potential relationship between L1 retotransposition and p53 loss or dysfunction. A significant increase (p=0.0013) in EGFP-positive cells was noted from 19.7% without amiR to 25.8% with amiR mp53 in HepG2, and an even more pronounced effect was observed in HT-1080 cells with intact TP53, where the percentage of EGFP-positive cells rose significantly (p=0. (0003) from 58.0% to 69.9% with amiR mp53. In TP53-deleted Saos2 cells, no significant difference was observed in EGFP-positive cells between the no amiR (3.7%) and amiR mp53 (3.5%) conditions (p=0.4576) (Fig. 14 B). This suggests that the amiR mp53 may target and inhibit TP53, leading to increased EGFP expression in TP53-intact cells, while having no effect where TP53 is absent.



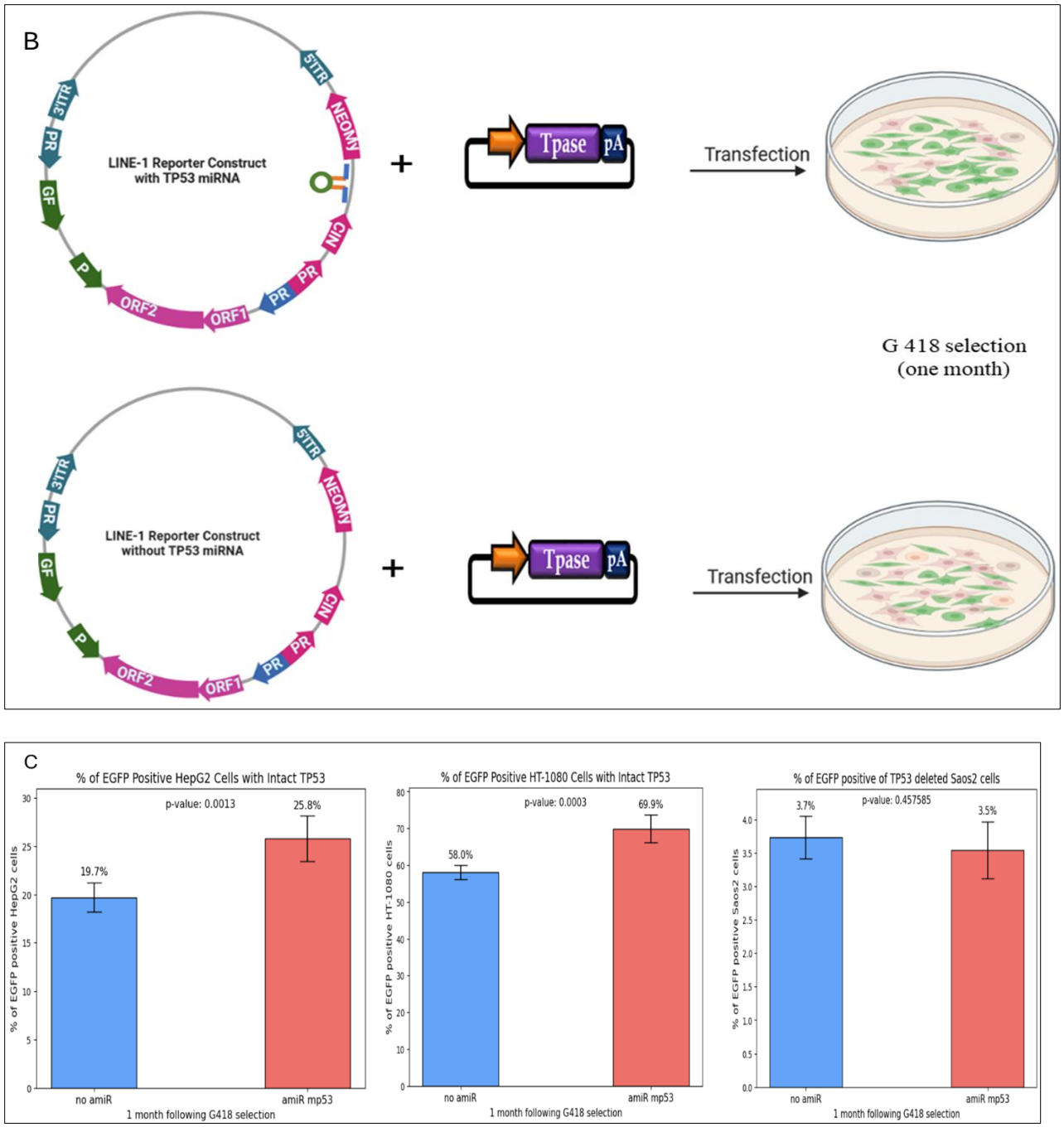


Figure 14. Silencing the TP53 gene associates with increased L1 retrotransposition. (A) Operation of our ORFeus reporter variant. (B) Overview of the *in vitro* assay system to study the effect of TP53 gene silencing on L1 retrotransposition. (C) Percentage of EGFP-positive cells in the amiR-expressing (blue bar) and non-expressing (red bar) cultures.

3.3 NEOADJUVANT TREATMENT INCREASES L1 EXPRESSION

3.3.1 Neoadjuvant treatment associated with increased ORF1p immunoreactivity.

It has been demonstrated that environmental stressors, including chemotherapeutic agents, UV light and gamma radiation, can induce aberrant L1 activation [84]. These findings have been corroborated *in vitro*. In this study, we sought to ascertain the relationship between intratumoral ORF1p positivity and the treatment history of the patients. The results demonstrated a notable elevation in L1 ORF1p expression among tumour specimens derived from patients who underwent neoadjuvant treatment, when compared to the untreated cohort. This observation suggests that antitumour therapy may potentially induce L1 expression. In the neoadjuvant therapy group, a significant proportion (p-value = 0.000001) of samples exhibited elevated ORF1p immunoreactivity, reaching approximately 80% (Fig. 15 A). In contrast, the untreated group displayed a higher proportion of samples with lower ORF1p intensities. To gain further insight into these findings, we conducted an additional analysis, testing ORF1p immunoreactivity across various cancer locations in patients who received neoadjuvant therapy (Fig 15 B) and those who were untreated(Fig 15). Our observations indicated that in patients who underwent neoadjuvant treatment, ORF1p immunoreactivity was generally elevated in cancers of the uterus, lung (SCLC), ovary, breast and pancreas, compared to untreated patients. This suggests a possible role of therapy in enhancing ORF1p expression. It is noteworthy that kidney cancer displayed lower ORF1p immunoreactivity, even in patients who received neoadjuvant therapy, in comparison to those who were untreated (Fig. 15 B&C). It may suggest that certain chemotherapeutic agents play significant role in inducing retrotransposition-mediated genome rearrangements, which could potentially lead to chemoresistance or the formation of a second primary cancer.

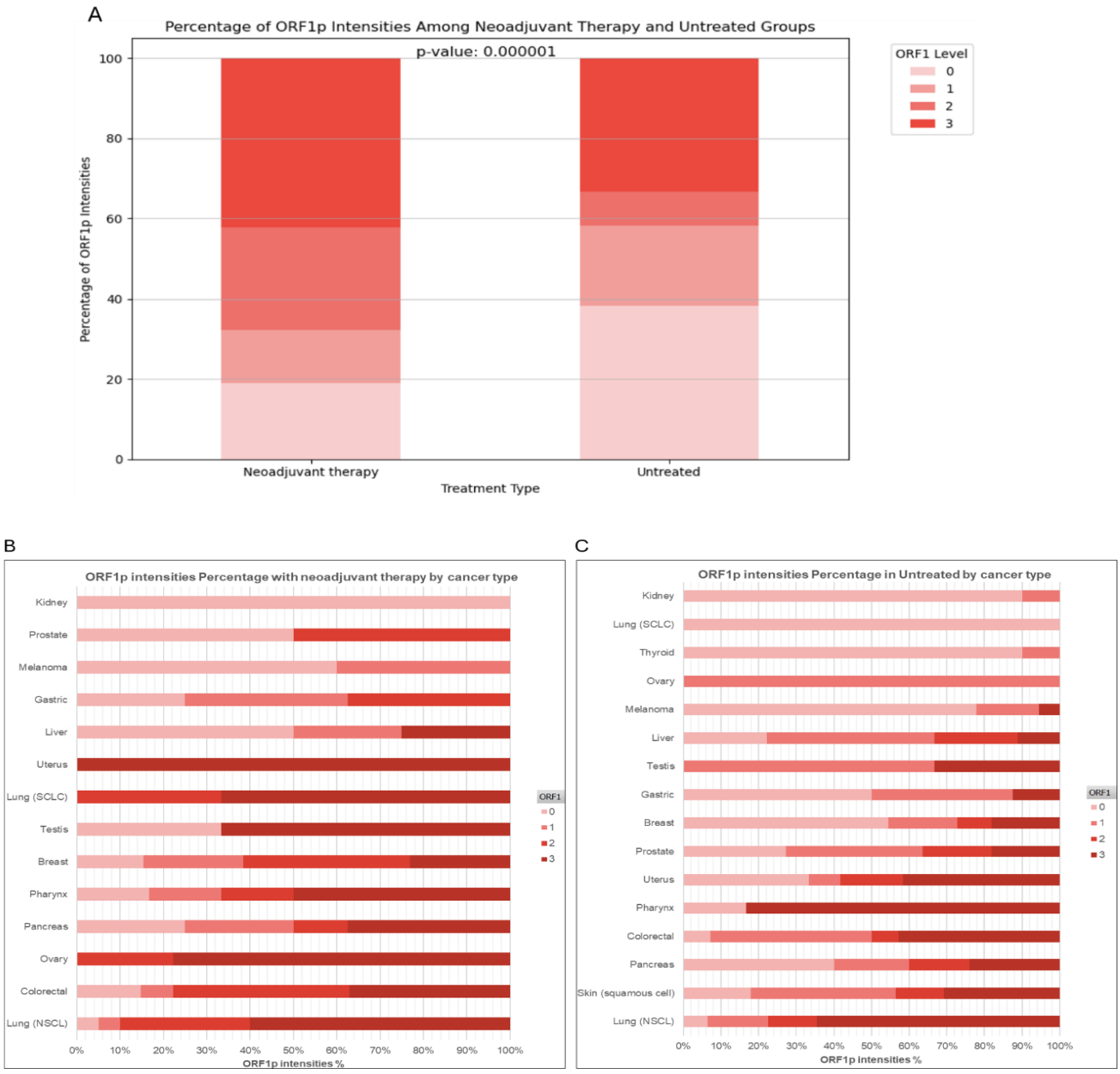


Figure 15. Neoadjuvant therapy is linked to higher levels of ORF1p immunoreactivity. The stacked bar graph shows the percentage of ORF1p intensities in neoadjuvant-treated and non-treated samples.

3.3.2 Paclitaxel increases the rate of L1 retrotransposition in vitro

To verify the results of the tumour bank study, namely that neoadjuvant therapy may result in an increase in ORF1p immunoreactivity, the technology described in Kopasz et al was employed [79]. HepG2 cells were transfected with a L1 reporter construct and the SB100 transposase. Following the complete selection with neomycin, one group was treated with 5nM paclitaxel for 72 hours,

while the other remained untreated. The number of EGFP-positive HepG2 cells was determined by FACS in both groups (Fig. 16). The data revealed that the paclitaxel-treated group exhibited 42.0% EGFP-positive cells, while the untreated group demonstrated 37.6% EGFP-positive cells. The observed difference is statistically significant, with a p-value of 0.0012. These findings offer insight into the potential impact of cancer treatments such as paclitaxel on L1 elements activity.

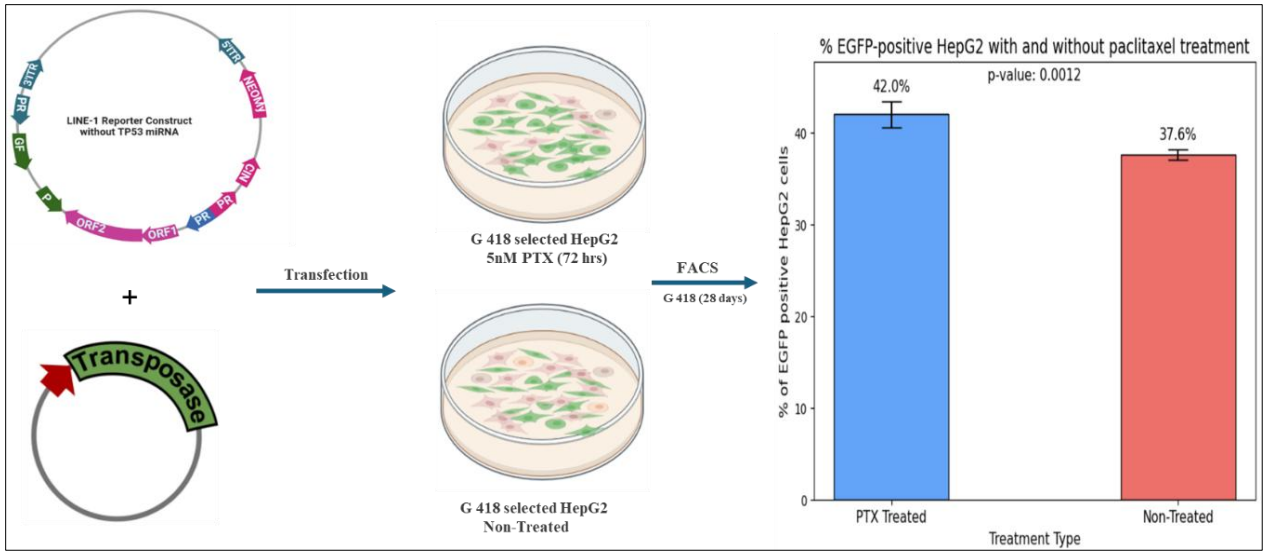


Figure 16. Flowchart alongside a bar graph that illustrates the effects of paclitaxel (PTX) treatment on HepG2 cells transfected with a L1 reporter construct. The graph on the right compares the percentage of EGFP-positive cells in the PTX-treated and non-treated groups.

3.4 L1 ORF1p IS A GOOD BIOMARKER FOR CERVICAL CANCER.

While evaluating the L1 ORF1p IHC staining results of tumour samples, a specific immunoreactivity pattern was observed in cervical cancers. Therefore, a more comprehensive investigation was conducted of additional cervical tissue samples, comprising a detailed analysis of 143 cervical specimens. A total of 20 cases of CIN I, 46 cases of CIN II, 14 cases of CIN III, 32 cases of invasive cancer, and 24 cases of non-dysplastic cervical epithelia (including seven cases of atrophy) were examined in detail. The expression of ORF1p, Ki67, and p16 was evaluated through immunohistochemical staining for entire 143 samples. P16 and Ki-67 are valuable diagnostic and assessment tools for the identification and evaluation of cervical neoplasia, particularly cervical intraepithelial neoplasia (CIN) and cervical cancer [85]. Their combination offers the potential for more accurate and informative results regarding the development of cervical

cancer. Therefore, we employed them in parallel with ORF1p immunostaining to demonstrate the promising use of ORF1p as a biomarker for cervical cancer.

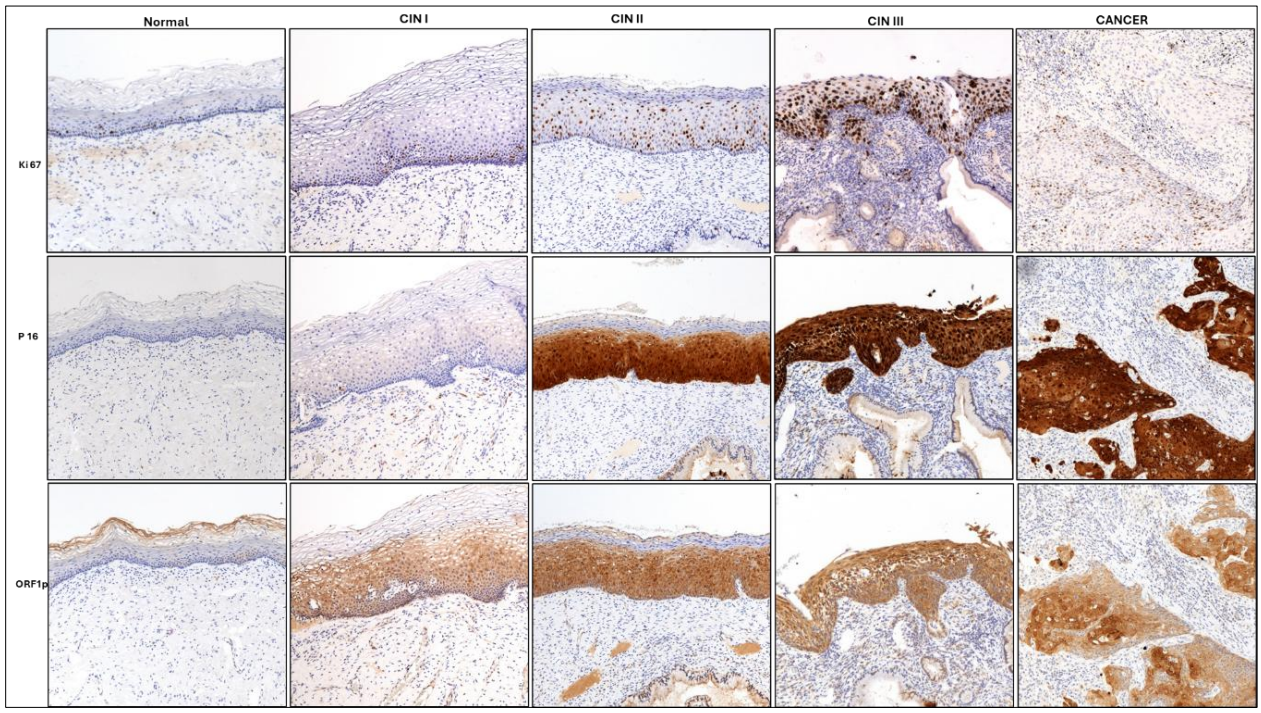


Figure 17. IHC staining of ORF1p, p16, and Ki67 of cervical samples representing various stages of cervical carcinogenesis, including normal, CIN I, II, and III, and cervical cancer, scale bar: 200 μm.

3.4.1 L1 ORF1p is expressed from early stage of cervical neoplasia.

To examine the expression pattern of OFR1p, normal cervical tissues and different grades of cervical cancer were stained with ORF1p, p16 and Ki67. We found that none of these markers are expressed in normal cervical tissue. However, as the tissue progresses to CIN I (Fig. 17), a slight increase in marker expression becomes evident, particularly in the lower third of the epithelium for Ki67 and P16, while ORF1p exhibits more pronounced expression in the upper two-thirds. The transition to CIN II is characterised by a substantial increase in all markers, with expression typically spanning the lower two-thirds of the epithelium (Fig. 17). This trend intensifies in CIN III, where Ki67, P16, and ORF1p all demonstrate strong, transepithelial expression throughout the entire thickness of the epithelium. In malignant tissue, all three markers display extensive and intense staining across the disorganized tissue architecture. The ORF1p diverged from the typical pattern observed in CIN I, demonstrating a rising expression with higher grades of CIN towards cancer. In contrast, it was barely or not observed in normal cervical epithelium (Fig. 17). The

expression of Ki67 diverged from the normal pattern in the CIN II stage only. Similarly, aberrant expression of the p16 marker is typically not detectable prior to the CIN II stage (Fig. 17). The representative images of CIN I at higher magnification clearly demonstrate patterns of Ki67 and p16 immunostaining that are characteristic of normal cervical epithelium. It is noteworthy that the Ki67 staining is positive in the basal layer of the epithelium, which is consistent with the established pattern of Ki67 expression in cervical tissue. This indicates that cell division is primarily occurring in the basal layer, while overexpression of ORF1p is readily discernible from normal staining patterns (Fig. 18). It is noteworthy that the ORF1p immunoreactivity does not extend to the full thickness of the cervical epithelia. The staining is most intense in the lower two-thirds of the epithelium, with a gradual decrease towards the surface (Fig. 18). This pattern suggests that the target protein is more highly expressed in the basal and parabasal layers, which are actively dividing. Considering these findings, we developed a scoring system which is a modified version of the Immunoreactivity Scale (IRS) by Remmele and Stegner [78]. This modified scoring system assigns a PP score ranging from 0 to 3, indicating the extent of epithelial involvement, from no staining (0) to transepithelial staining (3), which spans the entire thickness of the epithelium. Similarly, the SI score, also ranging from 0 to 3, quantifies the intensity of staining, from no colour reaction (0) to a strong colour reaction (3) (Table 2). The final IRS score is calculated using the formula 'IRS Score = PP x SI. This modified system provides a structured approach for assessing and categorising of ORF1p expression in cervical tissues, facilitating research and diagnostic purposes. From this point onwards, this modified scoring system was employed to assess the results of the ORF1p immunostaining in our entire sample panel. According to the new modified IRS scoring system, it was a notable increase in ORF1p immunoreactivity as the tissue progressed from normal to malignant state. As the tissue progresses to CIN1, there is a notable increase in ORF1p expression . The upward trend continues in the case of CIN2, The most significant increase is evident in the CIN3 and invasive cancer stages (Table 3). In general, the results demonstrated that ORF1p immunoreactivity was present in 98.2% of pre- and malignant cervical lesions (any score), and approximately 80% of normal cervical specimens did not exhibit-immunoreactive ORF1p (Fig. 19A). Of the 24 samples, only five displayed minimal ORF1p immunoreactivity in the basal one- or two-thirds of the epithelial thickness, with a score of 1–2. In the atrophic cases, only one sample exhibited minimal ORF1p immunoreactivity, with a score of 2 (Fig. 19A). Conversely, ORF1p immunoreactivity was identified in all cases of CIN I, with a

common localization to the lower third of the cervical epithelium, displaying a median score of 2 (Fig. 19A and Table 3). Similarly, ORF1p immunopositivity was observed in all CIN II cases, with a principal distribution across the lower two-thirds of the cervical epithelium and an elevated median score of 4 (Fig. 19A and Table 3). Also, all CIN III samples demonstrated ORF1p immunoreactivity, with a strong trans-epithelial positivity and a further high median score of 9. Of the 14 cases of CIN III, only one exhibited a moderate immunoreaction (score 4), while the remaining 13 samples showed a strong immunoreaction (score 6–9) (Fig. 19A). The samples from invasive cancer cases also exhibited strong ORF1p immunopositivity, with a median score of 9 (Fig. 19A and Table 3). Of the 32 cancer cases, 30 showed a strong immunoreaction (score 6–9) (Fig. 19A). The statistical analysis demonstrated a markedly heightened immunoreaction for ORF1p in dysplastic lesions (CIN I-III and cancer) when compared to normal tissue ($P < 0.0001$). Additionally, a markedly stronger immunoreaction for ORF1p was discerned in CIN I cases when compared to normal cases ($P < 0.0001$) (Fig. 19A). Furthermore, the immunoreactivity scores were found to be significantly higher in HSILs (high-grade squamous cell lesions, including CIN II and CIN III) compared to the CIN I group ($P = 0.001$). This suggests that HSILs are at a considerably elevated risk of progressing to invasive cervical cancer.

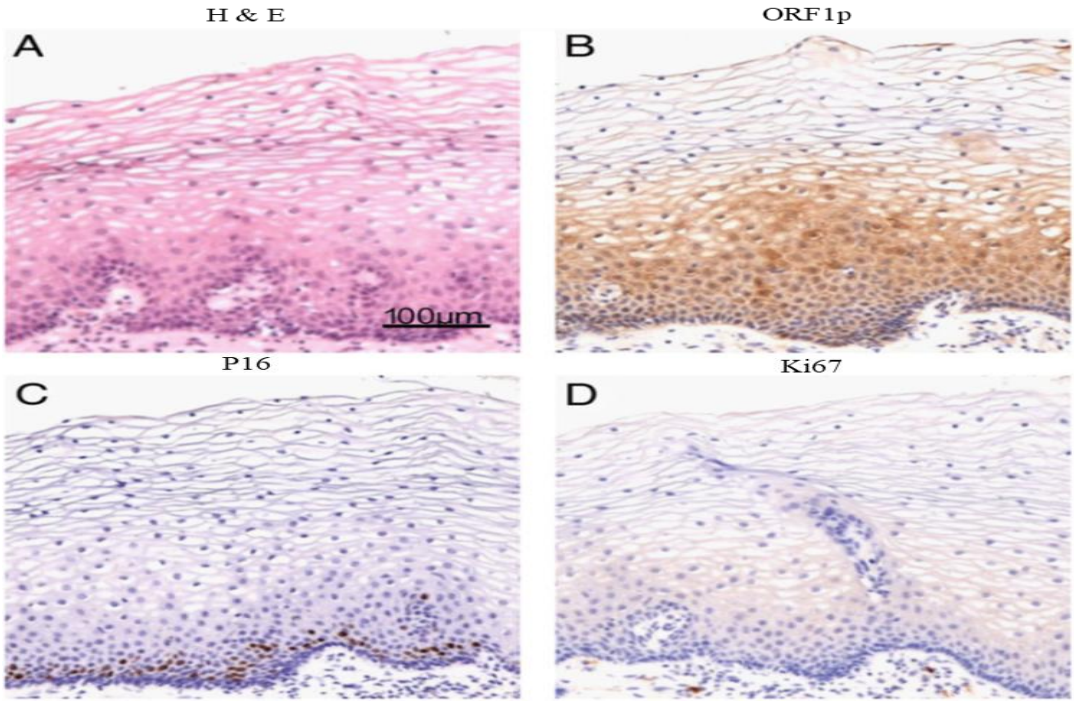


Figure 18. (A) HE staining, (B-D) ORF1p, p16, and Ki67 IHC on CIN I samples. Scale bar: 100 µm.

A			
Positively stained epithelial thickness (PP)	**PP score**	**Intensity of staining (SI)**	**SI score**
No staining	0	No color reaction	0
Basal one-third	1	Mild reaction	1
Basal two-thirds	2	Moderate reaction	2
Transepithelial	3	Strong reaction	3
B			
IRS Scores (PP × SI)		**IRS Classification**	
0		Negative	
1–2		Positive, mild immunoreaction	
3–4		Positive, moderate immunoreaction	
6–9		Positive, strong immunoreaction	

Table 2. A modified version of the immunoreactive scale (IRS) scoring system for the evaluation of ORF1p IHC in cervical specimens.

			ORF1p immunoreaction scores			
	Normal	**Atrophy**	**CIN1**	**CIN2**	**CIN3**	**Invasive cancer**
Sample number	24	7	20	46	14	32
Median	0.000	0.000	2.000	4.000	9.000	9.000
Mean	0.375	0.286	2.300	3.870	8.000	7.781
SD	0.770	0.756	1.174	1.809	1.710	2.393

Table 3. An overview of ORF1p immunoreaction scores in cervical tissue specimens, categorised according to the different stages of CIN and invasive cancer. n= 143.

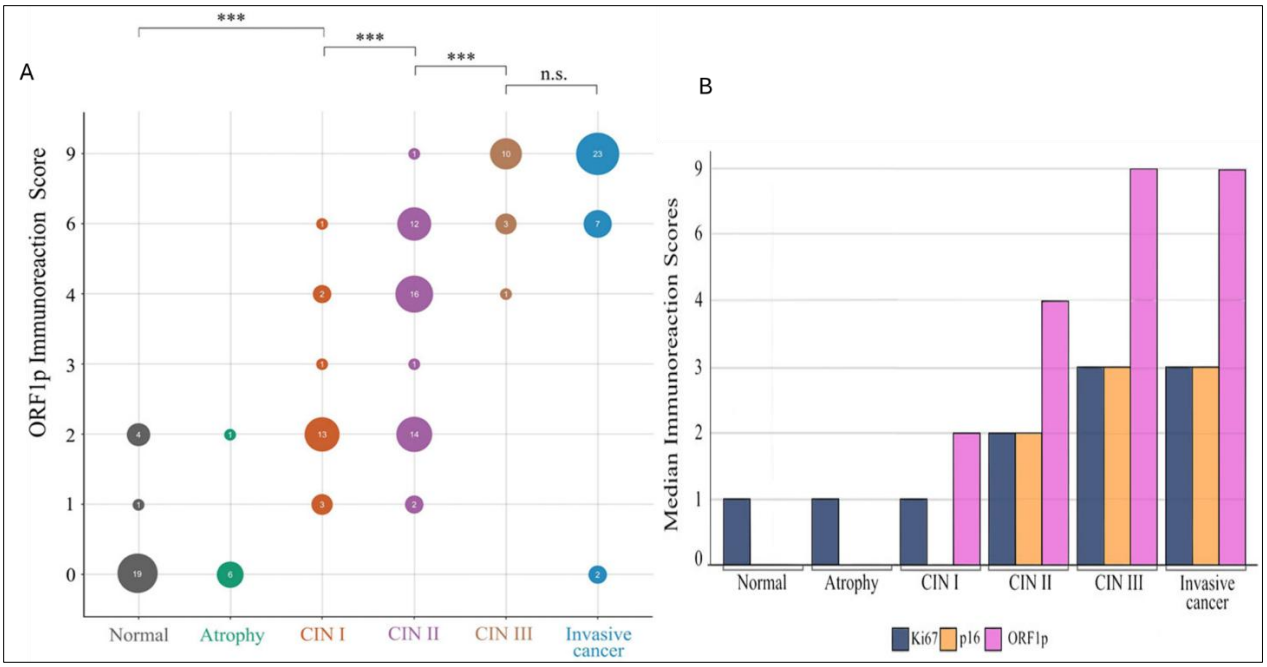


Figure 19. ORF1p expression increases as cervical tissue changes from normal to cancerous. (A) ORF1p immunoreaction scores across cervical tissue samples with different stages. The size and number in each bubble show the number of individuals with a particular score in each sample. (B) Comparison of the median immunoreaction scores of Ki67, p16, and ORF1p across different stages of cervical tissue. Three stars indicate $P < 0.001$: n.s.= not significant, n = 143.

Next, the expression pattern of Ki 67, P 16 and ORF1p in all cervical epithelium samples was compared. It was found that in normal and atrophic tissues, only Ki67 exhibited a low level of expression (score of 1), whereas p16 and ORF1p were not detected (Fig. 19B). As the tissue progresses to CIN I the level of Ki67 expression remains at 1, while ORF1p begins to show expression with a score of 2 (Fig. 19B). It is of interest to note that p16 remains undetected throughout this process. In cases of CIN II, all three markers are observed to exhibit elevated expression levels. Both Ki67 and p16 attain a score of 2, while ORF1p reaches 4. This trend intensifies in CIN III, with Ki67 and p16 increasing further to 3 and ORF1p rising significantly to 9 (Fig. 19B). In invasive cancer, the pattern remains like that observed in CIN III, with Ki67 and p16 at 3 and ORF1p at 9. This suggests that ORF1p expression commenced at an early stage (CIN I) and then reached its peak at CIN III, maintaining this level in invasive cancer (Fig. 19B). This pattern suggests that these biomarkers, particularly ORF1p, may prove beneficial in the detection and staging of cervical neoplasia. ORF1p may prove to be the most effective early marker for cervical cancer and its progression. Next, the results of the p16 and Ki67 immunoreactions were

evaluated for the whole sample panel using a three-grade scale to quantify the degree of immunoreactions[70, 86]. The results demonstrated that the intensities of p16 and Ki67 immunoreactions were significantly higher in the group comprising all neoplastic and dysplastic lesions (CIN I-III and cancer) compared to the control group ($P < 0.001$) (Fig . 20 A&B). The data demonstrate a general upward trend in both mean and median staining scores as the tissue abnormality increases, with a particularly sharp rise between CIN1 and CIN2. The highest p16 and ki 67 staining scores are observed in specimens classified as CIN3 and invasive cancer. It is noteworthy that the mean and median values are in close alignment in the majority of categories, indicating a symmetric distribution of staining scores (Fig. 20 A&B).

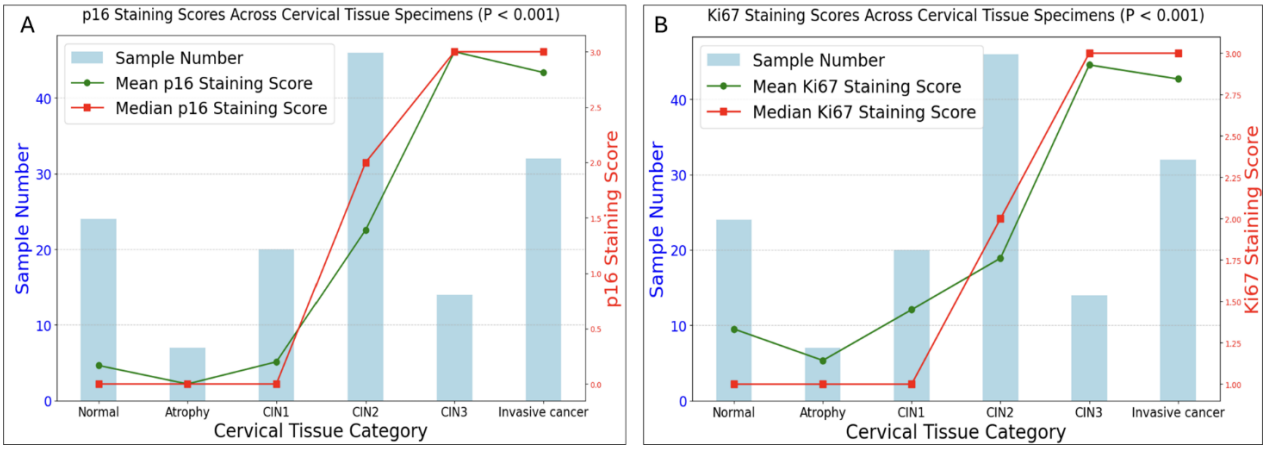


Figure 20. P16 and Ki67 are more immunoreactive in CIN I-III and cancer than in normal cervical tissues.

3.4.2 ORF1p staining is a reliable test to diagnose cervical cancer cases

To evaluate the accuracy of ORF1p IHC as a diagnostic test for neoplastic and dysplastic cervical cancer, we assessed the diagnostic test performance parameters. The calculated positive predictive value (PPV) and negative predictive value (NPV) of ORF1p staining which were 77% and 100%, respectively, when compared to non-dysplastic samples, indicating a high degree of accuracy in the detection of CIN I. The test result exhibits an exceptionally high level of sensitivity, with a narrow confidence interval approaching 100%. This finding is consistent with the absence of false negatives observed in the confusion matrix (Fig. 21A). The specificity is also satisfactory at approximately 80%, although with a wider confidence interval, which reflects the presence of some false positives. The positive predictive value is approximately 75-80%, indicating that a positive test result is relatively reliable but not definitive. The negative predictive value is approximately 100%, which is consistent with the absence of false negatives and provides

further evidence that a negative test result is highly reliable for ruling out CIN I. The results demonstrate that ORF1p staining is an effective and highly sensitive test for CINI, with excellent specificity for identifying positive cases and particularly strong negative predictive value (Fig. 21A). Furthermore, the calculated positive predictive value (PPV) and negative predictive value (NPV) of ORF1p staining were 83% and 94%, respectively, when invasive cancer was compared to non-dysplastic cases, indicating a high degree of accuracy in the detection of cervical cancer cases. While the positive and negative test results are highly reliable, they are not definitive for ruling out invasive cancer (Fig. 21A&B).

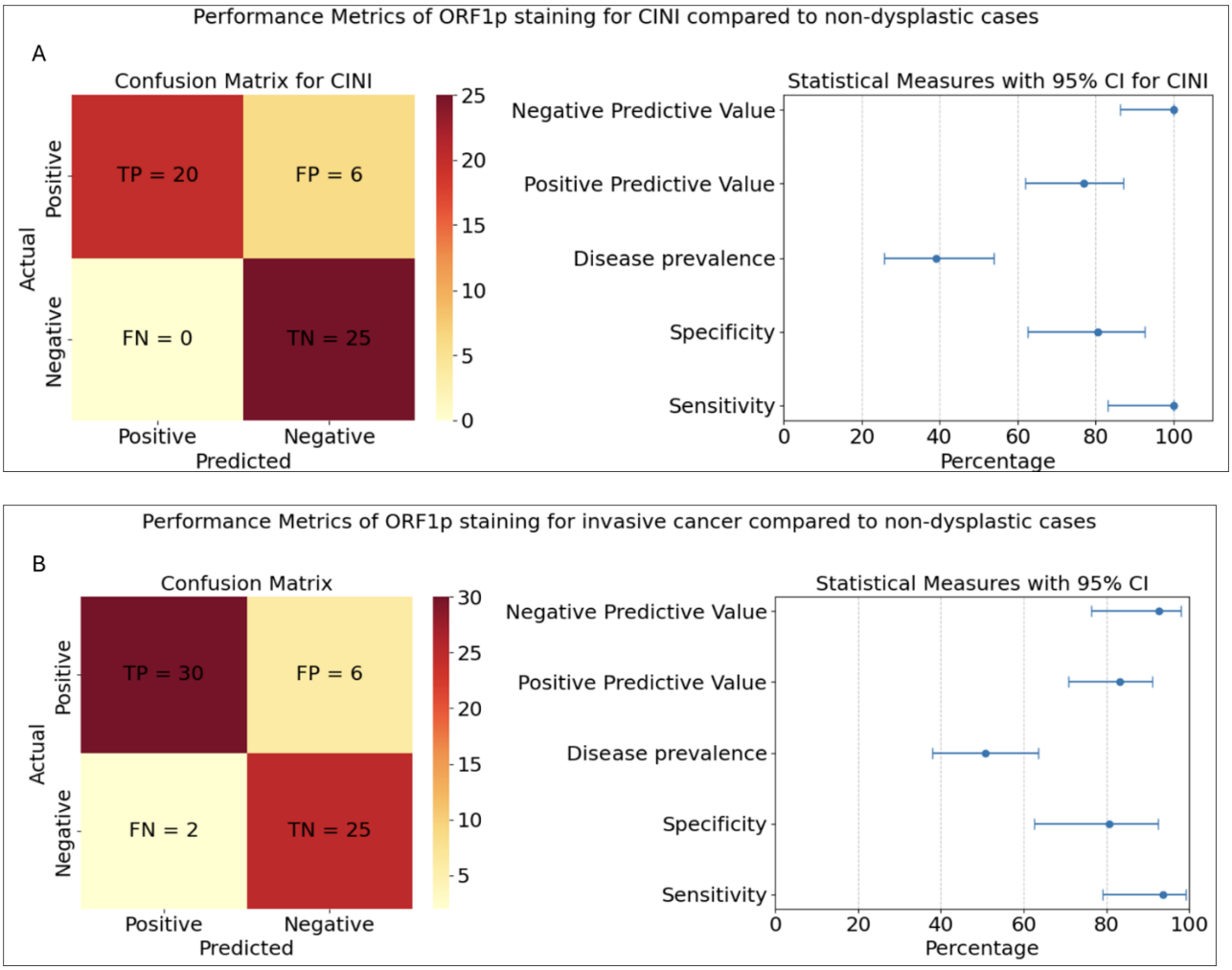


Figure 21. ORF1p staining is an effective method for identifying positive cases of cervical cancer. The confusion matrix (on the right) provides a detailed comparison between the actual CIN1 and invasive cases and the predicted results based on ORF1p staining. On the left, the forest plot displays five statistical

measures, each represented with a 95% confidence interval (CI), to provide further insight into the performance of the ORF1p staining.

4. DISCUSSION

The findings of this thesis offered new insights into the expression patterns and potential functions of the L1 ORF1 protein (ORF1p) in a wide range of human somatic tissues and cancers. Our findings indicate that L1 ORF1p is commonly absent or expressed at low levels in the majority of normal somatic tissues. Furthermore, our cohort demonstrated a limited expression of ORF1p in normal somatic tissues, which is consistent with previous reports indicating that L1 retrotransposition is tightly regulated and largely silenced in differentiated cells [87, 88]. The low expression level observed in striated muscles and the epidermis is also consistent with previous literature indicating the presence of residual L1 activity in specific cell types [89]. The tissue-specific regulation of ORF1p indicates that its expression may be associated with cellular differentiation and proliferation states. The absence or low expression of ORF1p in normal tissues may be attributed to various mechanisms, including epigenetic silencing, transcriptional regulation, and post-transcriptional control, as described previously in introduction of this thesis [10, 90]. The expression of L1 ORF1p has been identified as a distinctive feature of various cancers by multiple research groups [54, 91]. Specifically, it has been demonstrated that L1 upregulation occurs at an early stage of development of certain tumour types, such as ovarian cancer, as evidenced by the accumulation of ORF1p in precursor lesions of the fallopian tube [92]. In line with these findings, we detected an increase in ORF1p expression intensity of immunostaining with the stages of tumourigenesis, which suggests that retrotransposition plays a significant role in tumour progression [93, 94]. The frequent overexpression of ORF1p in a wide range of human malignancies highlights the potential deregulation of L1 activity during malignant transformation. The observed variability in ORF1p expression levels across different cancer types indicates the presence of complex, context-dependent mechanisms governing L1 regulation in various malignant settings. The high ORF1p expression in skin basal cell carcinoma, cervical cancer, oesophageal cancer, and non-small cell lung cancer may suggest a more prominent role for L1 activation in the pathogenesis of these cancers [54]. To further investigate the potential link between ORF1p expression and tumour progression, we assessed the correlation between ORF1p immunoreactivity, clinical stage and grade. The findings of our investigation suggest a correlation between the proportion of ORF1 overexpression and both stage and grade. The positive correlation between ORF1p expression and advanced cancer stage and grade which mean that increased L1

activity may be associated with more aggressive tumour behaviour and poor clinical outcomes. This finding is in agree with previous studies that have linked elevated L1 retrotransposition to genomic instability and tumour progression [95–97]. During the screening process, a correlation was observed between ORF1p expression and cellular anaplasia, irrespective of the tumour type under consideration. Cellular anaplasia is a well-established phenotypic marker of dedifferentiation of a tumour cell, evaluated during histopathological examination according to a degree of grade. To examine this, we employed three grade categories: low, intermediate, and high histological undifferentiation grades. These grades were evaluated by an experienced histopathologist using haematoxylin-eosin staining. By avoiding the use of the tumour-type-specific grading system, the data can be homogenised and used on types. Furthermore, we emphasised the issue of intratumoral heterogeneity in L1 ORF1p expression, noting considerable variability within the same tumour samples. Some cancers have demonstrated a mosaic pattern of ORF1p expression, whereas others have exhibited a central-peripheral distribution. The intratumoral heterogeneity in ORF1p expression observed in this study represents a significant finding, as it emphasises the dynamic and context-dependent nature of L1 regulation within the tumour microenvironment. This heterogeneity may be driven by a variety of factors, including differential regulation of L1 ORF1p retrotransposon in distinct cellular subpopulations, the impact of local environmental clues, and the clonal evolution of the tumour [54, 98]. Our findings suggest that intratumorally heterogeneity of ORF1p expression may serve as a marker for precise grading and assessment of certain tumour types in clinical practice. In this regard, if ORF1p can be used as a biomarker for any cancer, the results of this thesis are particularly noteworthy regarding the potential of ORF1p as a biomarker for cervical neoplasia and cancer. the result provided comprehensive insight into the distinction between the expression patterns of ORF1p, p16, and Ki67 markers in cervical dysplasia and cancer. The results demonstrate that ORF1p expression increases in a progressive manner with the development of cervical dysplasia and the transition to invasive cancer. This suggests that ORF1p has the potential to serve as a valuable early marker, capable of differentiating between various grades of cervical intraepithelial neoplasia (CIN) through the presentation of distinct immunoreaction patterns. This is particularly noteworthy considering the limitations of current screening methods, such as Pap smears and HPV testing, which have been demonstrated to miss a significant number of precancerous lesions and early-stage cancers [75, 99]. Also, ORF1p immunostaining suggests a new accurate and trustworthy

method for recognizing between higher CIN grades than the traditionally used p16 and Ki67 markers. Though p16 immunostaining has been shown to be effective in identifying high-risk HPV-linked lesions, its analysis can sometimes be biased and prone to interobserver changeability [100]. This is proven by the findings of Clark et al. [101], who published that 19% of the assessed HSIL samples had to be recategorize as LSIL because positive p16 immunostainings were misunderstood at the time of the original diagnosis. The accompanying presence of a mild ORF1p immunoreaction characteristic of LSIL, when combined with ORF1p immunostaining, may aid the pathologist in avoiding such a mistake due to the presence of positive p16 immunostaining. Here we reported that a mild or strong ORF1p immunoreaction is a more usual finding in HSIL cases. In comparison to the other biomarkers, the distinct practical clinical value of ORF1p staining test is likely to be its ability to differentiate normal cervical epithelium from CIN I lesions with a high degree of trustworthiness. This is confirmed by the mean ORF1p immunoreaction score of 0.37 for normal epithelium and 2.30 for CIN I lesions. The 100% negative predictive value is specifically markable, given that all cases of CIN I showed a positive ORF1p immunoreaction. The PPV and NPV values were likewise evaluated for other sample groups, which keep further evidence of the valuableness of ORF1p staining test. The high sensitivity and negative predictive value of ORF1p staining for cervical intraepithelial neoplasia and invasive cancer provide further evidence to support its potential as a supplementary diagnostic tool in cervical cancer screening and management. The utilisation of ORF1p as a potential biomarker for recognising CIN lesions is analogous to the implement paradigm wherein ORF1p expression has been used as a diagnostic marker for HPV-independent differentiated vulvar intraepithelial neoplasia (dVIN), the precursor lesion of vulval squamous cell carcinoma [102]. The rare ORF1p immunopositivity in normal cervical epithelium was confirmed by the recent study by McKerrow et al, who found that L1 mRNA expression is rarely detectable in normal epithelial cells across different tissue types [103]. This suggests that L1 expression is not entirely confined to epithelial cells and that leaky expression may occur. In agreement with our findings, many authors, including Lavasanifar et al., Taylor et al., Ponomaryova et al. and Kou et al., have proposed that L1 shows considerable promise as both a biomarker and a therapeutic target in cancer. Its expression and hypomethylation are associated with a range of cancers, making it a valuable tool for diagnostics and prognostics. Furthermore, the detection of L1 proteins in the bloodstream provides an effective, non-invasive method for cancer monitoring [56, 104–106]. These findings provide a comprehensive insight into

the differential expression patterns of ORF1p, p16, and Ki67 markers in cervical dysplasia and cancer. ORF1p has been identified as a promising early marker, capable of distinguishing between various grades of cervical intraepithelial neoplasia (CIN) by offering distinct staining patterns. To examine the impact of specific intracellular regulatory mechanisms that control the L1 ORF1p retrotransposon in distinct cellular subpopulations, we investigated the relationship between ORF1p expression patterns and the status of cellular tumour suppressor genes and oncogenes. Our findings indicated a moderate positive correlation between L1 ORF1p expression and TP53 immunoreactivity, suggesting that tumours with dysfunctional TP53 are more susceptible to elevated L1 activity. This observation is consistent with the findings of previous studies [38, 107]. The observed correlation between ORF1p and TP53 aberrations in our cohort suggests that the loss of p53 function may contribute to the activation of L1 elements, potentially through the disruption of pathways that normally suppress retrotransposon activity. The results of our *in vitro* experiment provide corroboration for this hypothesis, demonstrating that TP53 silencing with artificial microRNA resulted in increased L1 retrotransposition in cell lines with intact p53 function. Similarly, a significant correlation was identified between ORF1p expression and mutations in BRCA1/2 and HER2, which are known to be involved in DNA repair and cell growth regulation, respectively [108, 109]. These findings indicate that the interplay between L1 activity and oncogenic pathways may play a role in the development and progression of certain cancers. The role of L1 retrotransposition in carcinogenesis can be elucidated by the observation that inhibition of L1 has been shown to inhibit tumour initiation and progression in preclinical experiments using reverse transcriptase inhibitors (RTIs). It has been demonstrated that the administration of reverse transcriptase inhibitors (RTIs) can result in a reduction in L1 activity and a decrease in genomic instability in cancer cells. In clinical trials, reverse transcriptase inhibitors (RTIs) have been investigated as a monotherapy or in combination with other cancer treatments, including chemotherapy and immune checkpoint inhibitors [110]. For example, a histone deacetylase (HDAC) inhibitor, ivaltinostat, when administered in combination with gemcitabine and erlotinib, resulted in an improvement in progression-free survival in patients with pancreatic adenocarcinoma [111]. These findings underscore the significance of retrotransposition in cancer progression and highlight the importance of further research into the particular tumour types that are most affected by retrotransposition-driven tumour progression, with the aim of improving the clinical outcomes of individuals afflicted with cancer. However, the relationship between L1

expression and neoadjuvant treatment in cancer patients remains poorly understood. In this study, we aimed to investigate whether L1 ORF1p expression is altered in cancer patients undergoing neoadjuvant therapy compared to those who did not receive treatment. One of objective of our research was to ascertain whether there is a difference in L1 ORF1p expression between cancer patients undergoing neoadjuvant therapy and those who did not receive treatment. The results demonstrated a notable elevation in L1 ORF1p expression among the cohort that received neoadjuvant therapy, indicating the potential for specific chemotherapeutic agents or irradiation to stimulate L1 retrotransposition. It is probable that the impact of chemotherapy on L1 retrotransposition is intricate and may fluctuate contingent on the specific chemotherapeutic agent employed, the type of cancer being treated, and other as yet unidentified factors. The observed increase in ORF1p levels in post-neoadjuvant treatment tumour samples, in comparison to those of untreated cases, suggests that certain cancer therapies may potentially lead to the activation of L1 retrotransposition. This finding is consistent with previous reports that have linked various environmental stressors, including chemotherapeutic agents and radiation, to the derepression of L1 elements [112, 113]. The discovery that neoadjuvant therapy, particularly chemoradiotherapy, is associated with elevated L1 ORF1p expression highlighted the necessity to elucidate the influence of cancer treatments on retroelement activity. Our *in vitro* experiments with paclitaxel support the notion that specific chemotherapeutic agents can stimulate L1 activity, potentially resulting in augmented genomic instability and the emergence of resistant cancer clones. Finally, further research is needed and should focus on explaining the mechanisms by which chemotherapy and other environmental stressors activate L1, as well as the potential for combining RTIs with conventional cancer therapies to suppress L1 activity and improve patient outcomes. It is important to note that the measurements presented in this study are specific to the L1 ORF1 protein. In light of the numerous post-translational regulatory mechanisms involved in the control of L1, further investigation is required to ascertain whether the observed protein expression correlates with the occurrence of novel retrotransposition events. Furthermore, additional research should be conducted to elucidate the mechanisms by which chemotherapy and other environmental stressors activate L1. Furthermore, additional research should be conducted to elucidate the mechanisms by which chemotherapy and other environmental stressors activate L1. Additionally, the potential for combining RTIs with conventional cancer therapies to suppress L1 activity and improve patient outcomes should be investigated.

SUMMARY

Cancer is currently the second leading cause of death in the European Union, following cardiovascular diseases. In recent years, its incidence has risen, largely due to an aging population, which places a significant burden on healthcare systems and society. Several studies have shown that the development of cancer is directly linked to the activity of L1 (L1) elements, the only currently active mobile genetic elements in the human genome, which make up about 17% of it. While most L1 elements are inactive due to mutations, some remain active and may contribute to genomic instability, influence gene expression, and potentially lead to diseases, including cancer. This thesis investigates the expression pattern of the ORF1p protein, encoded by L1, in various human cancers and explores its potential as a diagnostic biomarker. Using immunohistochemistry on tissue microarrays, the study analyzed 590 samples from 21 different tumor types, as well as samples from cervical intraepithelial neoplasia (CIN) and normal tissues. The results revealed that L1-ORF1p is minimally expressed in most normal somatic tissues but is frequently expressed in a wide range of cancers, including cervical cancer, non-small cell lung cancer, esophageal cancer, and basal cell carcinoma of the skin. Elevated L1-ORF1p expression was notably associated with higher clinical stages and histological grades, indicating a link to cancer progression. Moreover, intratumoral heterogeneity in L1-ORF1p expression was observed, with higher expression in less differentiated tumor areas. The study found a significant correlation between L1-ORF1p overexpression and mutations in tumor suppressor genes, particularly TP53. Tumors with mutated TP53 exhibited higher levels of L1-ORF1p than those with wild-type TP53, suggesting that the loss of function of TP53 may promote L1 activation and contribute to oncogenesis. This relation was confirmed by our experiments in tissue culture, where we found that silencing endogenous TP53 with artificial microRNAs (amiRs) increased the number of L1 retrotransposition events. The findings suggest that L1-ORF1p is a novel biomarker for both cancer diagnosis and prognosis. Its expression patterns could help distinguish between low and high-grade lesions, potentially improving diagnostic accuracy and patient management. Additionally, targeting L1-ORF1p could offer new therapeutic options in cancer treatment. The study highlights the importance of further research to elucidate the mechanisms through which L1 activity contributes to tumorigenesis and to explore its potential clinical applications.

CONCLUSION

Our research findings indicate that L1 ORF1p exhibits high immunoreactivity in cancerous tissues when compared to their normal somatic tissues. It has been observed that the activity of L1 ORF1p is particularly elevated in cases of skin basal cell carcinoma, cervical cancer, oesophageal cancer and non-small cell lung cancer. Our research indicates that ORF1p expression is associated with cancer progression, with higher levels observed in more advanced tumours. This indicates that L1 may be involved in tumour progression and could potentially be used to predict the stage of advancement of a tumour. Our research indicates a correlation between ORF1p expression and TP53 overexpression, suggesting a potential link between L1 activity and p53-mediated tumour suppression pathways. The results of the in vitro experiments demonstrated that L1 retrotransposition increased when TP53 was silenced. Furthermore, our research indicated a correlation between neoadjuvant therapy and elevated ORF1p levels in multiple cancer types. This was confirmed by an *in vitro* assay indicating that certain cancer treatments (PTX) may stimulate retrotransposon activity. Our study identified L1 ORF1p as a potential biomarker for cervical cancer. It was expressed at the early stages of neoplasia, and the ORF1p score demonstrated efficacy in differentiating between normal cervical tissue and various grades of cervical intraepithelial neoplasia (CIN) and invasive cancer.

ACKNOWLEDGEMENTS

First of all, I am grateful to Almighty ALLAH, who gave me the strength and ability to accomplish this work successfully. Countless salutation upon our beloved HOLY PROPHET HAZRAT MUHAMMAD (PBUH). Secondly, with utmost gratification, I would like to express my profound and intense sense of indebtedness to my ever-affectionate worthy supervisor, Dr. Lajos Mátés, PhD, Biological research centre, University of Szeged, Hungary. His proficient counselling, trust, valuable suggestions, boundless forbearance, indefatigable help with anything, anywhere, anytime, consummate advice, thought provoking instructions and giving me enough independence to decide the things throughout the study, helped me to grow in both my competence and in confidence as a researcher. I would like to express my gratitude to Dr. Andrea- Mátés Nagy who helped me a lot regarding the data analysis and making the illustrations and for all her time she spent to teach me. I am unfathomable indebted to Prof. Dr. Dux László, MD, PhD, DSc, Head of the Multidisciplinary medical science doctoral scchool, for his encouragement, full support and cooperation during research work. I am also very much thankful for the insightful advice, technical support and continuous help from Dr. Farkas Sükösd, MD, PhD institute of pathology, Faculty of medicine, University of Szeged, Hungary during the whole period of study. I am grateful for the extended assistance and encouragement obtained from, Dr. Gabriella Pankotai-Bodó, Dr. András Blastyák, Ivett Csikós and Sarolta Bankó and all the staff working at institute of pathology, Faculty of medicine, University of Szeged, Hungary. I would like to thank my PhD colleagues Dr. Réka Karkas, Dr. Gergely Imre and May Raya for all their positive support and constructive companionship throughout my study. I do not have words at my command to express my gratitude and profound admiration to my affectionate parents and family members. They provided me with their continuous support and love during my entire life. Finally, I would like to mention about the Tempus Public Foundation, Hungary and Higher Education Commission of Yemen who awarded me with this prestigious scholarship ‘Stipendium Hungaricum Scholarship Program’ through which I accomplished my dream of higher studies.

REFERENCES

[1] Kazazian HJ. Mobile elements: drivers of genome evolution. *Science* 2004; 303: 1626–1632.

[2] Lander E, Linton L, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature* 2001; 409: 860–921.

[3] Burns K, Boeke J. Human transposon tectonics. *Cell* 2012; 149: 740–752.

[4] Chuong E, Elde N, Feschotte C. Regulatory activities of transposable elements: from conflicts to benefits. *Nat Rev Genet* 2017; 18: 71–86.

[5] Hancks D, Kazazian HJ. Roles for retrotransposon insertions in human disease. *Mob DNA* 2016; 7: 9.

[6] Deininger P, Moran J, Batzer M, et al. Mobile elements and mammalian genome evolution. *Curr Opin Genet Dev* 2021; 66: 12–20.

[7] Mita P, Boeke J. How retrotransposons shape genome regulation. *Curr Opin Genet Dev* 2016; 37: 90–100.

[8] Zhang Y, others. Transcriptionally active HERV-H retrotransposons demarcate topologically associating domains in human pluripotent stem cells. *Nat Genet* 2019; 51: 1380–1388.

[9] Servant G, others. The Nuclease Activity of Mammalian L1 Retrotransposon Proteins. *J Mol Biol* 2017; 429: 2769–2782.

[10] Protasova MS, Andreeva TV, Rogaev EI. Factors Regulating the Activity of L1 Retrotransposons. *Genes* 2021; 12: 1562.

[11] Denli A, Narvaiza I, Kerman B, et al. Primate-specific ORF0 contributes to retrotransposon-mediated diversity. *Cell* 2015; 163: 583–593.

[12] Speek M. Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. *Mol Cell Biol* 2001; 21: 1973–1985.

[13] Moran J, Holmes S, Naas T, et al. High frequency retrotransposition in cultured mammalian cells. *Cell* 1996; 87: 917–927.

[14] Szak S, Pickeral O, Makalowski W, et al. Molecular archeology of L1 insertions in the human genome. *Genome Biol* 2002; 3: research0052.

[15] Khazina E, Weichenrieder O. Non-LTR retrotransposons encode noncanonical RRM domains in their first open reading frame. *Proc Natl Acad Sci USA* 2009; 106: 731–736.

[16] Dombroski B, Mathias S, Nanthakumar E, et al. Isolation of an active human transposable element. *Science* 1994; 266: 1612–1615.

[17] Feng Q, Moran J, Kazazian HJ, et al. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* 1996; 87: 905–916.

[18] Mathias S, Scott A, Kazazian HJ, et al. Reverse transcriptase encoded by a human transposable element. *Science* 1991; 254: 1808–1810.

[19] Ivancevic A, others. Horizontal transfer of BovB and L1 retrotransposons in eukaryotes. *Genome Biol* 2018; 19: 85.

[20] Taylor M, others. Affinity proteomics reveals human host factors implicated in discrete stages of L1 retrotransposition. *Cell* 2013; 155: 1034–1048.

[21] Macia A, others. Engineered L1 retrotransposition in nondividing human neurons. *Genome Res* 2017; 27: 335–348.

[22] Monot C, others. The specificity and flexibility of l1 reverse transcription priming at imperfect T-tracts. *PLoS Genet* 2013; 9: e1003499.

[23] Sultana T, others. The landscape of L1 retrotransposons in the human genome is shaped by pre-insertion sequence biases and post-insertion selection. *Mol Cell* 2019; 74: 555–570.
[24] Zhao K, others. Structural insights into the RNA and DNA specificities of human L1 ORF1 protein. *Cell* 2021; 184: 2284–2297.
[25] Ardeljan D, others. Cell fitness screens reveal a conflict between L1 retrotransposition and DNA replication. *Nat Struct Mol Biol* 2020; 27: 168–178.
[26] Flasch D, others. Genome-wide de novo L1 Retrotransposition Connects Endonuclease Activity with Replication. *Cell* 2019; 177: 837–851.
[27] Liu N, others. Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators. *Nature* 2018; 553: 228–232.
[28] Van Meter M, Kashyap M, Rezazadeh S, et al. SIRT6 represses L1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age. *Nat Commun* 2014; 5: 1–10.
[29] Vazquez BN, Thackray JK, Simonet NG, et al. SIRT7 mediates L1 elements transcriptional repression and their association with the nuclear lamina. *Nucleic Acids Res* 2019; 47: 7870–7885.
[30] Idica A, Sevrioukov EA, Zisoulis DG, et al. MicroRNA miR-128 represses L1 (L1) retrotransposition by down-regulating the nuclear import factor TNPO1. J Biol Chem. 2017 Dec 15;292(50):20494-20508.
[31] Fung L, Guzman H, Sevrioukov E, et al. miR-128 Restriction of L1 (L1) Retrotransposition Is Dependent on Targeting hnRNPA1 mRNA. Int J Mol Sci. 2019 Apr 21;20(8):1955.
[32] Zhang A, Dong B, Doucet AJ, et al. RNase L restricts the mobility of engineered retrotransposons in cultured human cells. Nucleic Acids Res. 2014 Apr;42(6):3803-20.
[33] Orecchini E, Frassinelli L, Galardi S, et al. Post-transcriptional regulation of L1 retrotransposition by AID/APOBEC and ADAR deaminases. *Chromosome Res* 2018; 26: 45–59.
[34] Morrish TA, Gilbert N, Myers JS, et al. DNA repair mediated by endonuclease-independent L1 retrotransposition. *Nat Genet* 2002; 31: 159–165.
[35] Suzuki J, Yamaguchi K, Kajikawa M , et al. Genetic evidence that the non-homologous end-joining repair pathway is involved in LINE retrotransposition. PLoS Genet. 2009 Apr;5(4):e1000461.
[36] Mita P, Sun X, Fenyö D, et al. BRCA1 and S phase DNA repair pathways restrict L1 retrotransposition in human cells. *Nat Struct Mol Biol* 2020; 27: 179–191.
[37] Tiwari B, Jones AE, Abrams JM. Transposons, p53 and Genome Security. Trends Genet. 2018 Nov;34(11):846-855.
[38] Wylie A, Jones A, D'Brot A, et al. p53 genes function to restrain mobile elements. *Genes Dev* 2016; 30: 64–77.
[39] Kawano K, Doucet AJ, Ueno M, et al. HIV-1 Vpr and p21 restrict L1 mobility. *Nucleic Acids Res* 2018; 46: 8454–8470.
[40] Sun X, Wang X, Tang Z, et al. Transcription factor profiling reveals molecular choreography and key regulators of human retrotransposon expression. Proc Natl Acad Sci U S A. 2018 Jun 12;115(24):E5526-E5535.
[41] Kazazian HJ, others. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature* 1988; 332: 164–166.

[42] Musova Z, others. A novel insertion of a rearranged L1 element in exon 44 of the dystrophin gene: Further evidence for possible bias in retroposon integration. *Biochem Biophys Res Commun* 2006; 347: 145–149.
[43] Suarez N, others. The Alteration of the C-terminal Region of Human L1 ORF1p Impairs L1 Retrotransposition and May Be Associated with Neurological Disorders. *Mol Neurobiol* 2020; 57: 1392–1405.
[44] Muotri A, others. L1 retrotransposition in neurons is modulated by MeCP2. *Nature* 2010; 468: 443–446.
[45] Crow M, others. Type I Interferons in Autoimmune Disease. *Annu Rev Immunol* 2019; 37: 247–267.
[46] Liu Y, others. Expression of p16INK4a in peripheral blood T-cells is a biomarker of human aging. *Aging Cell* 2009; 8: 439–448.
[47] De Cecco M, others. L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature* 2019; 566: 73–78.
[48] Rodriguez-Martin B, others. Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by L1 retrotransposition. *Nat Genet* 2020; 52: 306–319.
[49] Scott E, others. A hot L1 retrotransposon evades somatic repression and initiates human colorectal cancer. *Genome Res* 2016; 26: 745–755.
[50] Gasior S, others. The human L1 retrotransposon creates DNA double-strand breaks. *J Mol Biol* 2006; 357: 1383–1393.
[51] Deniz Ö, others. Endogenous retroviruses are a source of enhancers with oncogenic potential in acute myeloid leukaemia. *Nat Commun* 2020; 11: 3506.
[52] Chuong E, others. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science* 2016; 351: 1083–1087.
[53] Jang H, others. Transposable elements drive widespread expression of oncogenes in human cancers. *Nat Genet* 2019; 51: 611–617.
[54] Rodić N, Sharma R, Sharma R, et al. Long interspersed element-1 protein expression is a hallmark of many human cancers. *Am J Pathol* 2014; 184: 1280–1286.
[55] Ardeljan D, Taylor M, Ting D, et al. The human long interspersed element-1 retrotransposon: an emerging biomarker of neoplasia. *Clin Chem* 2017; 63: 816–822.
[56] Taylor M, others. Ultrasensitive detection of circulating L1 ORF1p as a specific multi-cancer biomarker. *Cancer Discov* 2023; 13: 2532–2547.
[57] Bieging K, Mello S, Attardi L. Unravelling mechanisms of p53-mediated tumour suppression. *Nat Rev Cancer* 2014; 14: 359–370.
[58] Lane D, Levine A. p53 Research: the past thirty years and the next thirty years. *Cold Spring Harb Perspect Biol* 2010; 2: a000893.
[59] Montoya-Durango D, Liu Y, Teneng I, et al. Epigenetic control of mammalian L1 retrotransposon by retinoblastoma proteins. *Mutat Res* 2009; 665: 20–28.
[60] Chan C, others. Human papillomavirus infection and cervical cancer: epidemiology, screening, and vaccination-review of current perspectives. *J Oncol* 2019; 2019: 3257939.
[61] Sung H, Ferlay J, Siegel R, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021; 71: 209–249.
[62] Schiffman M, Wentzensen N, Wacholder S, et al. Human papillomavirus testing in the prevention of cervical cancer. *J Natl Cancer Inst* 2021; 103: 368–383.

[63] Wentzensen N, Schiffman M, Palmer T, et al. Triage of HPV positive women in cervical cancer screening. *J Clin Virol* 2020; 76: S49–S55.
[64] Stoler M, Schiffman M. Atypical squamous cells of undetermined significance-low-grade squamous intraepithelial lesion triage study. *JAMA* 2001; 285: 1500–1505.
[65] Creagh T, Bridger J, Kupek E, et al. Pathologist variation in reporting cervical borderline epithelial abnormalities and cervical intraepithelial neoplasia. *J Clin Pathol* 1995; 48: 59–60.
[66] Carreon J, Sherman M, Guillén D, et al. CIN2 is a much less reproducible and less valid diagnosis than CIN3: Results from a histological review of population-based cervical samples. *Int J Gynecol Pathol* 2007; 26: 441–446.
[67] Gage J, Schiffman M, Hunt W, et al. Cervical histopathology variability among laboratories: A population-based statewide investigation. *Am J Clin Pathol* 2013; 139: 330–335.
[68] Silva D, Gonçalves A, Cobucci R, et al. Immunohistochemical expression of p16, Ki-67 and p53 in cervical lesions - A systematic review. *Pathol Res Pract* 2017; 213: 723–729.
[69] Bullwinkel J, Baron-Lühr B, Lüdemann A, et al. Ki-67 protein is associated with ribosomal RNA transcription in quiescent and proliferating cells. *J Cell Physiol* 2006; 206: 624–635.
[70] Nam E, others. Expression of the p16 and Ki-67 in relation to the grade of cervical intraepithelial neoplasia and high-risk human papillomavirus infection. *J Gynecol Oncol* 2008; 19: 162–168.
[71] Ordi J, Garcia S, del Pino M, et al. p16 INK4a immunostaining identifies occult CIN lesions in HPV-positive women. *Int J Gynecol Pathol* 2009; 28: 90–97.
[72] O'Neill C, McCluggage W. p16 expression in the female genital tract and its value in diagnosis. *Adv Anat Pathol* 2006; 13: 8–15.
[73] Mills A, Paquette C, Castle P, et al. Risk stratification by p16 immunostaining of CIN1 biopsies: a retrospective study of patients from the quadrivalent HPV vaccine trials. *Am J Surg Pathol* 2015; 39: 611–617.
[74] Reuschenbach M, Wentzensen N, Dijkstra M, et al. p16INK4a immunohistochemistry in cervical biopsy specimens: A systematic review and meta-analysis of the interobserver agreement. *Am J Clin Pathol* 2014; 142: 767–772.
[75] Wentzensen N, Fetterman B, Castle P, et al. p16/Ki-67 Dual Stain Cytology for Detection of Cervical Precancer in HPV-Positive Women. *J Natl Cancer Inst* 2015; 107: djv257.
[76] Li X, Luo D, Zhang L, et al. Accurate interpretation of p53 immunohistochemical patterns is a surrogate biomarker for TP53 alterations in large B-cell lymphoma. *BMC Cancer* 2023; 23: 1–11.
[77] Köbel M, Ronnett BM, Singh N, et al. Interpretation of P53 Immunohistochemistry in Endometrial Carcinomas: Toward Increased Reproducibility. *Int J Gynecol Pathol* 2019; 38: S123–S131.
[78] Remmele W, Stegner H. Recommendation for uniform definition of an immunoreactive score (IRS) for immunohistochemical estrogen receptor detection (ER-ICA) in breast cancer tissue. *Pathologe* 1987; 8: 138–140.
[79] Kopasz A, Pusztai D, Karkas R, et al. A versatile transposon-based technology to generate loss- and gain-of-function phenotypes in the mouse liver. *BMC Biol* 2022; 20: 74.

[80] An W, Dai L, Niewiadomska AM, et al. Characterization of a synthetic human L1 retrotransposon ORFeus-Hs. *Mob DNA* 2011; 2: 2.
[81] Ergün S, Buschmann C, Heukeshoven J, et al. Cell type-specific expression of L1 open reading frames 1 and 2 in fetal and adult human tissues. *J Biol Chem* 2004; 279: 27753–27763.
[82] Lüönd F, Tiede S, Christofori G. Breast cancer as an example of tumour heterogeneity and tumour cell plasticity during malignant progression. *Br J Cancer* 2021; 125: 164–175.
[83] Mátés L, Chuah M, Belay E, et al. Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. *Nat Genet* 2009; 41: 753–761.
[84] Terasaki N, Goodier JL, Cheung LE, et al. In vitro screening for compounds that enhance human L1 mobilization. *PloS One* 2013; 8: e74629.
[85] Harper D, Anderson R, Baker E, et al. Cost-effectiveness of p16/Ki-67 Dual-Stained Cytology Reflex Following Co-testing with hrHPV Genotyping for Cervical Cancer Screening. *Cancer Prev Res Phila* 2023; 16: 393–404.
[86] Alshenawy H. Evaluation of p16, human papillomavirus capsid protein L1 and Ki-67 in cervical intraepithelial lesions: Potential utility in diagnosis and prognosis. *Pathol Res Pr* 2014; 210: 916–921.
[87] Hagan C, Sheffield R, Zerler B. Histone H3.3 and H2A.Z antagonize the transcriptional activity of the L1 retrotransposon. *Proc Natl Acad Sci U A* 2003; 100: 12294–12299.
[88] García-Pérez J, Marchetto M, Muotri A, et al. L1 retrotransposition in human embryonic stem cells. *Hum Mol Genet* 2007; 16: 1569–1577.
[89] Belancio VP, Roy-Engel AM, Pochampally RR, et al. Somatic expression of L1 elements in human tissues. *Nucleic Acids Res* 2010; 38: 3909–3922.
[90] Rosser JM, An W. L1 expression and regulation in humans and rodents. *Front Biosci Elite Ed* 2012; 4: 2203.
[91] Ardeljan D, Wang X, Oghbaie M, et al. L1 ORF2p expression is nearly imperceptible in human cancers. *Mob DNA* 2020; 11: 1.
[92] Pisanic T 2nd, Asaka S, Lin S, et al. Long interspersed nuclear element 1 Retrotransposons become deregulated during the development of ovarian Cancer precursor lesions. *Am J Pathol* 2019; 189: 513–520.
[93] Zolfaghari MA, Karimi A, Kalantari E, et al. A comparative study of long interspersed element-1 protein immunoreactivity in cutaneous malignancies. *BMC Cancer* 2020; 20: 1–10.
[94] Hur K, Cejas P, Feliu J, et al. Hypomethylation of long interspersed nuclear element-1 (L1) leads to activation of proto-oncogenes in human colorectal cancer metastasis. *Gut* 2014; 63: 635–646.
[95] Rodić N, Burns K. Long Interspersed Element-1 (L1): Passenger or Driver in Human Neoplasms? *PLoS Genet* 2013; 9: e1003402.
[96] Miki Y, Nishisho I, Horii A, et al. Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Res* 1992; 52: 643–645.
[97] Schauer S, Carreira P, Shukla R, et al. L1 retrotransposition is a common feature of mammalian hepatocarcinogenesis. *Genome Res* 2018; 28: 639–653.
[98] Nguyen T, Carreira P, Sanchez-Luque F, et al. L1 retrotransposon heterogeneity in ovarian tumor cell evolution. *Cell Rep* 2018; 23: 3730–3740.

[99] Nayar R, Wilbur D. The Pap Test and Bethesda 2014. 'The reports of my demise have been greatly exaggerated.' (Mark Twain). *Acta Cytol* 2015; 59: 121–132.
[100] van Bogaert L. P16INK4a immunocytochemistry/immunohistochemistry: Need for scoring uniformization to be clinically useful in gynecological pathology. *Ann Diagn Pathol* 2012; 16: 422–426.
[101] Clark J, others. Overdiagnosis of HSIL on cervical biopsy: Errors in p16 immunohistochemistry implementation. *Hum Pathol* 2016; 55: 51–56.
[102] Hofstetter G, Mildner M, Tschandl P, et al. ORF1p Is a Potential Novel Diagnostic Marker for Differentiated Vulvar Intraepithelial Neoplasia. *Int J Gynecol Pathol* 2023; 42: 201–206.
[103] McKerrow W, others. L1 retrotransposon expression in cancerous, epithelial and neuronal cells revealed by 5' single-cell RNA-Seq. *Nucleic Acids Res* 2023; 51: 2033–2045.
[104] Lavasanifar A, Sharp C, Korte E, et al. Long interspersed nuclear element-1 mobilization as a target in cancer diagnostics, prognostics and therapeutics. *Clin Chim Acta* 2019; 493: 52–62.
[105] Ponomaryova A, Rykova E, Gervas P, et al. Aberrant Methylation of L1 Transposable Elements: A Search for Cancer Biomarkers. *Cells* 2020; 9: 2017.
[106] Kou Y, Cen S, Li X. Research and application on L1 in diagnosis and treatment of tumorigenesis. *Yi Chuan* 2021; 43: 571–579.
[107] Levine A, Ting D, Greenbaum B. P53 and the defenses against genome instability caused by transposons and repetitive elements. *Bioessays* 2016; 38: 508–513.
[108] Chenais B. L1 retrotransposons and their impact on genome stability and cancer in humans. *FEBS Open Bio* 2020; 10: 2167–2181.
[109] Choi J, Hwang S, Ahn K. Interplay between BRCA1 and transposable elements in the genome. *Genomics Inf* 2018; 16: e21.
[110] Rajurkar M, others. Reverse Transcriptase Inhibition Disrupts Repeat Element Life Cycle in Colorectal Cancer. *Cancer Discov* 2022; 12: 1462–1481.
[111] Jo J, Jung D, Lee H, et al. A phase I/II study of ivaltinostat combined with gemcitabine and erlotinib in patients with untreated locally advanced or metastatic pancreatic adenocarcinoma. *Int J Cancer* 2022; 151: 1565–1577.
[112] Giorgi G, Marcantonio P, Del Re B. L1 retrotransposition in human neuroblastoma cells is affected by oxidative stress. *Cell Tissue Res* 2011; 346: 383–391.
[113] Terasaki N, Goodier J, Cheung L, et al. In vitro screening for compounds that enhance human L1 mobilization. *PLoS One* 2013; 8: e74629.